

Needless quarrel over research

The governments of the United States and Japan are deadlocked in their negotiation of a new agreement on basic research collaboration. Most of the faults seem to lie with the United States.

If a researcher takes time off from his regular job to work at a laboratory elsewhere, who in chauvinistic terms can be judged to gain? The country at which the researcher chooses to spend his time, or that to which he eventually returns? The simplest answer is that there is no simple answer, and that the balance of advantage will depend on the circumstances. Yet it now appears that an attempt to define a general rule on just this question is one of the issues over which the governments of the United States and Japan are deadlocked in their attempt to renegotiate their agreement on cooperation in science and technology. Some fear that the negotiations will fail, others that they may succeed and, in the process, sour relations between Japan and the United States. There is also a risk that a new agreement emerging from the protracted negotiations will damage the pattern of research in the United States, or Japan, or both. What circumstances have led two powerful governments to create such an impasse for themselves?

The original agreement on US-Japanese collaboration was signed during the Carter presidency, and should have been renewed last year. Instead, there have so far been four rounds of abortive negotiations, the most recent last month in Washington. Now, a US delegation is about to pack its bags for another visit to Tokyo, while the original agreement has been extended by a further two months. Neither side will say much about the reasons why the negotiations have run into such heavy weather, although there is little doubt that the proceedings are now deeply coloured by the growing conviction of the US government that Japan has won unfair advantages over many years from the openness of the basic research system in the United States.

Thus the United States is seeking to make the new agreement conform with principles that have no part in the existing agreement. It is asking for some kind of equality in the numbers of Japanese visiting US laboratories and those travelling in the other direction, arrangements on intellectual property rights allowing for an equitable sharing of the economic benefits of the research and an assurance that practical applications with military applications will not be transferred to hostile third parties. There is also some bickering about the way in which the new agreement on collaboration should be managed. To be fair to the United States, not all of this should be regarded as a direct consequence of US indignation about the imbalance of trade between the United States and Japan. The US government seems also to be genuinely anxious to negotiate an agreement that will serve as a model for other agreements on collaboration.

Philosophy

The underlying philosophy of the US position seems to be the supposition that even basic research has now become an essential ingredient of economic power and also both an economic cost (that of doing it) and an economic prize (the value of the applications eventually spawned). This seems to have been the thinking behind the paper presented by the United States to the ministerial meeting of the Organisation for Economic Cooperation and Development (OECD) last autumn, when the US government argued that the cost of basic research should be shared equally, by some yardstick such as gross national prod-

uct, among the industrial nations with access to the free or almost free markets in which the fruits of the research are traded. The principle, which sounds equitable, is seductive. But it is based on a misconception of the function of basic research, and in particular on the assumption that the sole purpose of what is spent on research is the same both inside and outside development laboratories — that of creating tangible wealth.

Relationships

If the United States sticks to this simple doctrine, it will sour relationships not merely with Japan but with others of its partners. The conduct of basic research, in the support of which the United States has a recent admirable record, has several quite different functions. Sometimes, as with most of astronomy and planetary exploration, the objective is to discover new understanding and knowledge for which, at the time, nobody can see practical applications; there may of course be side benefits, the training of students for example, but work of this kind is pure cost, repaid in coin of a cultural character. Much the same is true of expensive high-energy physics, although the spin-off represented by advances in, say, power engineering and data processing is far from negligible. Similarly, it is more than probable that as much as a half of the budget of the US National Institutes of Health is not supported for the sake of calculable economic or medical benefits, but in the general belief that a deeper understanding of living things cannot but eventually yield opportunities for medical advance. It is an act of faith, not an economic investment. It is also true, of course, that some basic research is much nearer to the market; this might, for example, have been the case with high-temperature superconductivity a year ago, although the failure so far to understand the properties of these materials in a simple way has diminished that chance.

It is plainly folly to treat all these cases as if they were the same for the purposes of the economic equitability of a bilateral agreement on research collaboration. Who would seriously suggest that a Japanese astronomer working on an accelerator experiment in the United States is somehow winning benefits but is somehow smart enough not to contribute anything in return to the success of the common enterprise? Much the same is true of the biomedical research splendidly supported in the United States. Naturally, there are also circumstances in which research experience in a foreign laboratory could be of direct economic benefit to the visitor's eventual employer but that is not apparently intended even under the renewed agreement. The United States is apparently talking of research projects at university and public laboratories, including unclassified research at national laboratories; Japan, recognizing that much basic research in Japan is done in company laboratories, is prepared to lean on companies to open some such laboratories to visitors from the United States.

Even so, it seems, people insist on counting heads. There appear to have been ten times as many Japanese researchers visiting the United States under the aegis of the existing agreement as there are Americans travelling in the opposite direction. None of that is particularly surprising, given the language barrier, the natural unwillingness in the competitive environment of



US research for people to disappear from the grant-winning scene for months or even years at a time and the increasing sense of self-completeness in the United States. The Japanese are rightly angered now to be blamed that the trade imbalance is mirrored in the movement of people. To everybody's credit, there is no official suggestion, whatever hotheaded congressmen may say, that the flow of Japanese will be reduced until that from the United States catches up; levelling up seems to be the goal. But the counting of heads, and the implication that the imbalance is somehow Japan's fault, makes very little sense.

The issue of intellectual property rights is another recipe for counting angels on the heads of pins. The United States is asking for some mechanism by which economic benefits arising from basic research projects defined as collaborations under the agreement should be shared. The Japanese side appears ready to concede the principle but is understandably unwilling to change its law on patents to conform with US practice. It is also, and again understandably, unable to concede the notion that contributions to basic research which are not themselves patentable, but from which patentable inventions later flow, should somehow qualify for compensation. This ambition seems to reflect an unreal appreciation on the US side of the relationship between basic research and development. Before the US negotiators, drawn largely from the State Department and the White House, take off for Tokyo next week, they might usefully consult more widely in Washington about the purposes of US spending on basic research. Even if the long-term goal is innovation or (last year's key-word), "competitiveness", the vehicle will be the harvest of shared understanding that flows from it, and the skill of those who work in the programme and who afterwards contribute to the health of US industry.

This is why the US stance is more than merely a threat to its relationships with Japan. What would happen to the quality of the US research effort if the notion underlying the negotiations with Japan took hold in the domestic grant-making agencies? If basic research is to be regarded as a tradeable economic commodity, many now-productive researchers will find themselves without support. There is also great danger if it is intended, as the OECD paper would suggest, to extend the principles which appear to inform the negotiations with Japan to other bilateral agreements on scientific and technical collaboration. By importing novel and unrealistic demands into the negotiations with Japan, the United States has made itself seem a bully. Less powerful countries than Japan would be even more resentful. Yet the plain truth is that, for the United States, the whole argument is a proxy for the other bilateral imbalances with Japan, especially in trade. Why not solve that directly, without risking turning a friendly nation into a less friendly one? □

Whose own goal?

Princess Anne should have spoken more carefully at last week's conference on AIDS.

THE conference on AIDS (acquired immune deficiency syndrome) in London last week should have helped engender among governments a general sense of the importance urgently of doing what can be done to inhibit the spread of the disease, which was a large part of the objective (see p. 377). By drawing into a national discussion of the dangers of doing nothing many of the African government's whose people are among the most seriously affected, the conference will also have helped to overcome the defensive reticence — no doubt born of helplessness — which has concealed the magnitude of the African problem from all but the most perceptive in the past few years. That the best inhibitor of the spread of AIDS is acknowledged to be public education is a measure of the difficulty that we all face: for the time being, medicine has no prophylaxis to offer while the known or prospective drug treatments are neither generally available nor more than palliatives. It is a depressing business

that the best remedy is public education, for the whole world knows how uncertain that is. The British government, which creditably played host to last week's conference, has recognized this truth for at least the past two years, during which it has mounted one of the most open publicity campaigns in the field.

It is also creditable that Princess Anne (now also called the Princess Royal), who is well-known for public-spirited good works such as being president of the Save The Children Fund and chancellor of the University of London, should have lent last week's occasion publicity and prestige by making a keynote speech at the opening of the conference. That is the good news. The bad is that Princess Anne's address was marred by the declaration that AIDS is mankind's "own goal" — a sporting metaphor implying self-inflicted damage. The implication was that male homosexuality, and sexual promiscuity in general, are somehow responsible for AIDS epidemic. At it happens, almost at the same time Dr Robert Gallo was giving a conference to Tokyo (see p. 401) a very different and less emotive answer to the proper question "Why now?". The dangers inherent in Princess Anne's remark are well illustrated by the way in which the general newspapers have leapt to the conclusion that the AIDS conference shows that "the permissive society" is over.

Sexual guilt, an oedipal phenomenon in freudian language, is widespread, accounting for much of the traditional reticence about sexual behaviour in advanced societies. In many countries of this kind, Britain for example, male homosexuality was until comparatively recently a criminal offence (as it remains if one of the partners is younger than 21). But there is no evidence that the decriminalizing legislation has increased the frequency of homosexual acts, although it has allowed male (and female) homosexuals to declare themselves as such, making them more recognizable and thus more conspicuous. It is the same with what is called "the permissive society", a euphemism for the notion that heterosexual intercourse may be nearly as common between unmarried as married adults. There is literally no worthwhile evidence to show that modern societies' willingness to acknowledge that patterns of behaviour often differ from social ideals has itself stimulated a change of behaviour that makes us all more vulnerable to AIDS. Even the increasing proportion of children born to unmarried women in most societies is as likely to be a consequence of changing relationships between the sexes as of greater promiscuity. The British government has implicitly acknowledged as much by emphasizing "safe sex" and the use of condoms in its most recent propaganda, as have the condom companies now allowed to advertise on television. It would be an even more daunting task to urge abstinence on people at large.

The plain truth, in any case, is that none of this is relevant to the emergence of the AIDS virus and its spread. The pattern of infection in many countries in Africa, the probable origin of the virus, suggests that male homosexuality is less relevant than heterosexual prostitution. Gallo's point last week was that it betokens a kind of complacency that, flushed with the success of the prophylaxis of most virus infections, people failed to anticipate the evolution of viruses specialized for the cells of the immune system. It is not, after all, that the phenomenon was unknown; infectious adult T-cell leukaemia, first recognized in Japan, is now known to be widely distributed and sometimes conspicuously so, as in the Caribbean. The AIDS virus HIV is necessarily much more insidious. Its chief victims so far have been male homosexuals, intravenous drug users and the recipients of contaminated blood products. That there will be some spread of AIDS in the heterosexual population is by now probably unavoidable, but the data on infectivity and incubation time which are now to hand are too meagre to sustain useful guesses. That matters little, for the outlook in many industrial countries is already grave enough to be a serious burden on society. The prospect for Africa could be calamitous. And the moral is that AIDS is not a moral failure but a viral infection. □

Global cooperation pledged after first AIDS summit

- One million cases by 1991 predicted
- Unanimous support for WHO programme

London

THE spirit of unity and consensus achieved at the first AIDS (acquired immune deficiency syndrome) summit of health ministers held in London last week was undeniable. Agreement was unanimous on the need for a global programme of health education, backed by sufficient resources, to stem the transmission of the human immunodeficiency virus (HIV). The 15-clause declaration that officially concluded the three-day meeting, attended by delegates from 149 countries, including 114 health ministers, pledged absolute commitment to the global AIDS strategy of the World Health Organisation (WHO), co-sponsor, with the British government, of the summit. How the meeting's laudable conclusions will be translated into political reality within individual countries is less certain.

The summit clearly achieved its stated objective to provide health ministers and senior policy-makers with a forum for discussing strategies for AIDS prevention and control with particular emphasis on information and education. The international publicity surrounding the summit will also have been welcomed.

Dr Jonathan Mann, director of WHO's AIDS programme, told the summit that between 5 and 10 million people are believed to be infected with HIV, with 75,000 fully developed cases of AIDS having been reported so far, although the true figure is likely to be nearer to 150,000. By 1991 it will rise to 1 million, said Mann.

Different modes of HIV transmission in different populations pose problems for global strategy. In Western Europe, North America and Australia, the virus

spreads mainly among homosexual and bisexual men and through intravenous drug abuse. In Africa and parts of the Caribbean, transmission is mainly through heterosexual contact.

The summit failed to make a recommendation on mandatory testing of individuals in high-risk groups. Attitudes towards compulsory testing vary greatly. According to Bulgaria's Professor Lyubomir Shindarov, HIV testing is compulsory for high-risk groups "such as blood donors, sperm and mother's milk donors, Bulgarian nationals coming back home after a long stay abroad, foreign students and workers who are going to spend more than a month in the country, pregnant women and newly married couples". Bulgaria has reported three cases of AIDS. The prevalence of the disease remains low in most communist countries. Dr Eugene Chazov, Soviet Minister of Health, told the summit that only one Soviet citizen is known to have been an AIDS victim, and only 33 cases of HIV infection have been reported. Of these, 18 had sexual contact with foreigners "who were probable sources of infection". Examination of 7,325 homosexuals and bisexuals and 43,486 individuals with multiple sexual contacts had not revealed a single seropositive case, Chazov said.

Professor She-Sheng Wang, of China, said that up to the end of last year, three cases of AIDS had been reported in China. Of those, two were foreign tourists and the third a former resident of Hong Kong who had lived in the United States.

Since 1985, a target population of 10,000 had been screened, revealing four seropositive haemophiliacs who had used imported factor VIII and seven seropositive foreigners. China, said Wang, would concentrate on preventing the import of AIDS. "Chinese law and traditional moral values prohibit homosexuality, sexual promiscuity and the abuse of drugs."

As of 12 January 1988, the Americas accounted for some 57,000 reported cases of AIDS (with 49,000 in the United States), Africa 8,693, Europe 8,775, Asia 224 and Oceania 742.

Undoubtedly the most important aspect of the summit was the exchange of information and ideas on different countries' handling of AIDS. Furthermore, the meeting demonstrated a greater willingness for many countries to concede that AIDS is a truly global issue, and that the threat it poses is real. **Simon Hadlington**

Novel Japanese superconductor

Tokyo

A GROUP of Japanese scientists at the Science and Technology Agency's National Research Institute for Metals in Tsukuba has reported a new high-temperature superconductor composed of bismuth-strontium-calcium-copper oxide. The development is unusual in that the new oxide does not contain any rare earth elements.

The new oxide, $\text{BiSrCaCu}_2\text{O}_x$, when annealed for several hours at just under 900°C , shows an onset in reduction of resistance at 120 K and zero resistance at 75 K. But there is a very sharp drop to almost zero resistance between 107 and 105 K. And Dr Hiroshi Maeda of the institute, whose group made the oxide, suspects that the new ceramic is composed of two phases, one with a critical temperature of 75 K and the other about 105 K.

Magnetic susceptibility measurements show clear evidence of the Meissner effect. And the oxide can be made to float in air when placed above a permanent magnet after soaking in liquid nitrogen. The oxide is black in colour but its crystal structure has not yet been determined.

All high T_c superconductors found so far contain rare earth elements, such as yttrium and lanthanum. In fact, Japan's Ministry of International Trade and Industry has been allotted ¥140 million (about \$1 million) in its fiscal 1988 budget to assess world supplies of rare earth elements because there were fears that Japan's supplies of the elements might dry up when the new superconductors reach the applications stage. Some theorists suggest, however, that it is the presence of Cu-O planes or chains within the atomic lattice that is crucial for superconductivity.

The results of the present discovery will be published in the *Japanese Journal of Applied Physics*. **David Swinbanks**

Messel pit saved

Bonn

A LAST-MINUTE decision by the West German *Land* of Hesse has prevented the fossil-rich Messel pit from being filled with rubbish. Preparations were nearly complete to turn the pit, located near Darmstadt, into a dump, despite objections from palaeontologists and environmentalists. The Messel pit is one of the richest known sources of Eocene fossils in the world.

The government of Hesse announced in mid-January its intention to seek another site for the dump, perhaps in a nearby forest. The decision must still be approved by local authorities. **Steven Dickman**

Shuttle launch date

Washington

THE National Aeronautics and Space Administration (NASA) has set 4 August as its target date for the resumption of space shuttle flights. The date may slip as the booster test programme continues, but NASA officials are promising a launch sometime in August unless new problems arise. The announcement was made by Admiral Richard H. Truly, NASA's associate administrator for space flight, before a congressional subcommittee on space science and applications. He said that some minor manufacturing flaws had been found after last December's booster test, but he insisted none was serious enough to threaten the launch date. **David Lindley**

CoCom trims embargo list but tightens controls

Paris & Munich

Representatives from the United States, Western Europe and Japan agreed on 28 January to shorten the list of items that may not be exported to the Soviet bloc, but to tighten controls on sensitive items, according to the motto "higher fences for fewer items".

The meeting of the Coordinating Committee for East-West Trade Policy (CoCom), an organization set up by NATO governments and Japan to monitor and limit East-West trade in goods and technology with potentially sensitive applications, took on a greater significance than usual because it reflects the current thaw in East-West relations. The meeting, at Versailles outside Paris, was nonetheless swaddled in secrecy and delegates from the 16 member states left without briefing the press.

The United States deputy secretary of state, John Whitehead, came to Paris to call for increased coordination between CoCom members in respecting existing restrictions and in sanctioning offenders.

Klaus Fuchs dies

London

Dr Klaus Fuchs, the nuclear physicist whose trial in 1950 for passing nuclear secrets to the Soviet Union became a turning point of the cold war, died last week in East Germany. Western comments on his death focused, inevitably, on his trial and nine years' imprisonment; the Soviet media did not mention it. The East German Party daily, *Neues Deutschland*, however, described it as a "political trial in connection with [Fuchs'] stance on the US monopoly of the nuclear bomb," and explained Fuchs' motive as "his steadfast opposition to the misuse of nuclear research findings for the implementation of imperialist power and intimidation politics".

Allowing for the difference in political vocabulary ("deterrence" on one side of the curtain may well appear "intimidation" on the other), this assessment is in fair agreement with those reached by the two Western biographies of Fuchs which have appeared in the past year (see *Nature* 327, 377; 1987, and 329, 774; 1987).

To *Neues Deutschland*, however, Fuchs' trial and sentence were merely one episode in his long life as "Communist, researcher and peace campaigner", putting its main emphasis on Fuchs' services to East Germany in the last 28 years of his life, as research physicist, initiator of East Germany's nuclear energy programme, and as a member of the Central Committee of the party.

Vera Rich

But European members, notably West Germany, have made no secret of their exasperation at the sheer number of items proscribed — thought to be as many as 300,000.

The computer and telecommunications industries in particular expect to increase their dealings with the East. Restrictions on computer memory size and speed were particularly troublesome.

One industry source complained that the CoCom rules were supposed only to hinder the export of "tomorrow's technology", not "yesterday's".

Eduard Shevardnadze, the Soviet foreign minister, is reported to echo this feeling, saying that the "cursed list" of secret technologies has already slowed development of benign domestic industries in the Soviet Union.

The US Congress is still smarting at the costly consequences of last year's breach of CoCom regulations by a Japanese and a Norwegian company. The Toshiba Machine Corporation and Kongsberg Vaapenfabrikk were found to have supplied machinery to the Soviet Union allowing the development of silent submarine motors. NATO governments have had to revise detection devices to counter this progress in Soviet knowhow.

But some in Congress are suspicious of the Pentagon's timing in releasing information about the Toshiba Machine case. They feel the Defense Department was trying to head off a pending congressional move to limit export restrictions. The current meeting may serve the purpose of restoring confidence in CoCom for members of Congress anxious for stricter controls, but at the same time heading off further liberalizing of laws covering export restrictions.

The threat of a deterioration in technology exchanges with the United States, if the Pentagon's anxieties were not soothed, has made European CoCom governments keen to show willing. It is therefore thought to be no coincidence that the French secret service chose the first day of the meeting to release details of the prosecution of six "techno-bandits" accused of selling sensitive electronic components to the Soviet Union.

An illegal supply network, involving a French company, Universal Testing Equipment, was uncovered by the French secret service in October 1987, but was kept low-key. Universal Testing Equipment allegedly provided a KGB attache in Paris with a catalogue of the West German company, Rohde and Schwartz, and subsequently supplied items from the catalogue to Moscow, using false shipping documents.

Soviets to launch German experiments

Munich

THE Soviet space programme received another boost last week with an agreement between the West German space hardware company Kayser Threde and the Soviet Union for three future launches of space experiments. The agreement reflects the dramatic improvement in the reputation of the Soviet programme coupled with the problems still plaguing the US shuttle programme.

Kayser Threde, which makes hardware for the West German Aerospace Research Establishment (DFVLR), has arranged to launch three payloads of 100–150 kg each on Soviet Cosmos rockets beginning in 1989. The rockets will carry both life science and materials science experiments that must be done under microgravity. The specific experiments have yet to be determined.

Kayser Threde had launched 25 experiments on the US shuttle, including some laser experiments on the West German D1 Spacelab mission. But its customers became impatient with the delay (which has just reached its two-year mark) in the shuttle programme resulting from the Challenger explosion. Kayser Threde's Reiner Klett said the delay was "bad for the customers and bad for our business".

Kayser Threde talked with both the People's Republic of China and the Soviet Union before deciding to sign up for Cosmos.

Negotiations had begun over a year ago and the agreement was actually reached in December, but it has come to the attention of the West only within the past week. The Soviets' price, Klett said, was competitive.

Steven Dickman

Although trade with the Soviet Union itself comprises less than one per cent of total exports, some West German industries expect to reap significant benefits from the agreement, especially in terms of sales to Hungary and other Soviet bloc countries. Many items remained on the CoCom list of restricted goods even though they were widely available on the world market.

One boon for West Germany lies in the decision to include specific descriptions of exceptions to the CoCom rules in the West German list. There is a simplified approval procedure for items included in the list of exceptions. The updated list is expected to be published some time this month.

But officials in West Germany have played down the economic significance of the meeting. The shortening of the list was seen primarily as a milestone in the political reconciliation between East and West.

Peter Coles & Steven Dickman

New Japanese observatory planned for detecting neutrinos

Tokyo

FLUSHED with the success of detecting neutrinos from a supernova, physicists at Tokyo University are proposing the construction of a gigantic underground observatory for neutrinos and nucleon decay that will dwarf their present facility. But a competing proposal from the same university to build a 1-GeV proton linear accelerator may stop the super-detector project in its tracks.

In February last year, a team headed by Professor Masatoshi Kishiba of Tokyo University detected a burst of neutrinos from the supernova 1987A in the Large Magellanic Cloud during the Kamiokande-II experiment being carried out 1 km underground in the Kamioka mine in Gifu Prefecture (*Nature* **326**, 121; 1987).

The Kamiokande observatory consists of a gigantic cylindrical water tank (15 m in diameter and 16 m deep) covered on the inside with highly sensitive photomultiplier tubes that detect Cerenkov radiation from high-energy electrons produced, for example, when neutrinos enter the tank.

The detector was originally built to observe proton decay — the grand unified theories predict that protons decay into a lepton plus meson(s) and that the leptons if charged will produce Cerenkov radiation that can be picked up by Kamiokande. But the half-life of protons has proved to be beyond the present facility's detection limit of about 10^{32} years.

The proposed 'Super-Kamiokande' detector will have a water capacity of 45,000 tons, enough to fill more than twenty Olympic swimming pools, and will be able to observe nucleon decay on a timescale of 10^{33} to 10^{34} years, as well as to detect solar and supernova neutrinos, and, perhaps, magnetic monopoles.

Construction of the super-detector will require excavation of a gigantic cave in Kamioka mine with a volume of around 10^5 m³ to house the cylindrical water tank which will be over 40 m high and 40 m in diameter. But this should not prove to be a major problem, according to one of the proponents of the scheme, because Mitsui Mining and Smelting Co. Ltd which owns Kamioka mine "is very cooperative to underground scientific research".

Cool crystal-clear water is readily available in the mine, and, after filtration, its average attenuation exceeds 35 m, more than enough to allow detection of bursts of Cerenkov radiation from the centre of the gigantic tank. And calculations by Dr T. Totsuka of Tokyo University indicate that Super-Kamiokande, when armed with 11,000 20-inch photomultiplier tubes and a computer, will be able positively to identify bursts of solar neutrinos by checking

that they are actually coming from the direction of the Sun.

Supporters of the scheme argue that it is becoming increasingly clear that the energies required to test the grand unified theories, which try to unify the strong, electromagnetic and weak forces, lie far beyond the capabilities of accelerators, whereas detection of proton decay with Super-Kamiokande would provide a unique opportunity to test the unification of the forces. But, ironically, it is Japan's accelerator physicists who may derail the

Super-Kamiokande project. The Institute of Nuclear Study at Tokyo University is proposing to build a 1-GeV proton linac to be set up next to the High Energy Physics Laboratory in Tsukuba which houses the world's most powerful electron-positron collider (see *Nature* **331**, 297; 1988).

Professor Akito Arima, vice-president of Tokyo University, says his university will have to choose later this year between the two proposals before requesting funds from the Ministry of Education, Science and Culture. He does not know which way the decision will go. But Super-Kamiokande has the advantage of being cheaper — ¥8,000 million (about \$60 million) as opposed to ¥40,000 million for the proton linac.

David Swinbanks

Greenwich Observatory moves a step closer to Cambridge

London

THE proposal to move the Royal Greenwich Observatory from its distinguished site at Herstmonceux Castle in Sussex to Cambridge overcame its final official hurdle last week when Cambridge City Council approved plans for the observatory's new £3-million building alongside the Institute of Astronomy. Work is expected to begin in September.

The decision by the Science and Engineering Research Council (SERC) to move the Royal Greenwich Observatory (RGO), announced some 18 months ago (see *Nature* **321**, 799; 1986), provoked considerable hostility, with SERC coming in for heavy criticism for failing to consult

for indigenous blue-collar workers. Furthermore, property prices are being pushed beyond the means of many of the local community. The University of Cambridge has undertaken to strive to accommodate the new RGO workforce.

The initial early bitterness within RGO caused by the announcement of SERC's intentions to move the observatory has now largely subsided, although staff are still not relishing the prospect of the inevitable upheaval. Furthermore, there will be a number of redundancies, probably confined to the lower grades of staff.

The remaining imponderable is how much SERC will be able to raise from the sale of the Herstmonceux site, which will be conducted under the watchful gaze of the country's conservationists, keen that the exquisite example of a mediaeval brick-built castle, a scheduled ancient monument and grade I listed building, remains accessible to the public (it is at present open from Easter to September, and attracts some 70,000 visitors a year).

SERC says the move will cost £5.5 million, and will pay for itself. Cheaper running and operating costs at its new site are expected to save RGO some £0.5 million a year for the first five years, with the sale of the Herstmonceux site making up the balance. The site will be officially placed on the market in the spring as three lots: castle, gardens, woodland and farmland; the office complex; and the equatorial group of six telescopes (currently forming part of an astronomy exhibition). Valuations are rumoured to run as high as £4 million, although the castle market is fickle and SERC will not be drawn on the matter of price speculation.

The satellite laser ranger will remain at the Herstmonceux site, with the likelihood that SERC will lease back from the buyer the land on which the ranger sits.

Simon Hadlington



adequately with the astronomy community. The move, originally scheduled for late 1989, now seems more likely to take place around the spring of 1990. Delays in SERC's timetable were caused when the initial planning application for the new building in Cambridge failed to provide sufficient information about how the move would affect local job opportunities and how the 100 RGO staff would be housed.

During the past decade, Cambridge has been the focus of a proliferation of high-technology business, dubbed the 'Cambridge phenomenon' and the city's administrators are becoming increasingly concerned that the influx of high-salaried executives is causing an imbalance in the local job market, with fewer opportunities

US commercial space companies prepare to fill shuttle vacuum

Washington

ALTHOUGH US space launches have decreased dramatically with the suspension of space shuttle flights, the Department of Transportation (DoT) is convinced that commercial launch services are about to take off. The reason for DoT's enthusiasm is an agreement signed last week between Martin Marietta and General Electric to launch 15 GE communications satellites aboard Martin Marietta's Titan rocket over the next several years.

The new agreement was made possible by a shift in US policy regarding commercial satellites that was formulated in 1983, but only made a reality after the accident to the Challenger space shuttle in 1986.

As originally conceived, the shuttle was to provide launch services to all comers, at a competitive price. But it became clear that a true cost accounting would make launching a satellite aboard the reusable shuttle considerably more expensive than aboard an expendable launch vehicle.

So, beginning in 1984, the Transportation Department was given responsibility for encouraging commercial companies to provide launch services without involving NASA (National Aeronautics and Space Administration) and the shuttle. But NASA remained very much in the commercial launch picture, as even with "total cost recovery", sending a communications satellite into orbit aboard the shuttle was approximately \$10 million cheaper than any commercial alternative.

But the Challenger accident changed all that. Companies seeking to launch satellites scrambled for launch vehicles. US rocket manufacturers were not prepared to take up the slack, so US companies explored the possibility of using Soviet, Chinese and French rockets.

In the summer of 1986, President Reagan announced that the shuttle was permanently out of the commercial launch business. That announcement made it clear that the future of commercial launches would lie with the private sector.

Richard E. Brackeen, president of Martin Marietta's commercial space launch subsidiary, calls the GE agreement a "benchmark order". Four additional satellites have also been committed to Titan launches. General Dynamics has signed agreements to launch three weather satellites for the National Oceanic and Atmospheric Administration aboard its Atlas rocket, as well as a Euro-

pean telecommunications satellite. McDonnell Douglas will also launch commercial payloads aboard its Delta rocket, including 2 satellites for British television.

In addition to the aerospace giants, some smaller companies are looking for a toehold in the competitive space launch

arena. Space Services, American Rocket Company and Conatec are all involved in development programmes.

The next stage in the 'privatization' of space will be clearer when the White House announces details of its new space policy signed by the president last month. In addition to calling for further exploration of Mars and the Moon, it gives a significant boost to the Industrial Space Facility, a far less expensive but more limited commercial version of the NASA space station.

Joseph Palca

Soviets pay a rare visit

Las Vegas

THE United States and the Soviet Union last week completed the "first practical steps" on a quest to verify superpower underground nuclear weapons tests when a 20-member Soviet team ended a five-day visit to the Nevada Test Site, 65 miles north-west of Las Vegas. The Soviet visit follows a visit by a US delegation earlier in January to the Soviet test site at Semipalatinsk.

Now that both sides have seen each other's remote, top-secret nuclear test sites, details of an agreement to explode up to a 150-kiloton nuclear weapon on both sides in order to experiment with measuring devices will come from negotiations in Geneva continuing on 15 February, according to a joint statement. Joint experiments could begin as soon as May.

Dr Robert Barker, the physicist who led the US delegation, said the two scientific teams "exchanged more than words". Soviet Foreign Minister Igor Palenikh, head of the Soviet team, said the process created more trust between the two countries, possibly leading to more agreements.

The joint experiments are intended to enable both sides to agree on improved verification measures necessary to ratify the US-Soviet Threshold Test Ban Treaty of 1974 and Peaceful Nuclear Explosions Treaty of 1976. At least one test on each site is planned.

Both sides agreed to study several ways to measure tests. The methods include US-supported CORTEX (Continuous Reflectometry for Radiation versus Time Experiment), plus the Soviet hydrodynamic and seismic models.

During the visit, the two delegations exchanged samples of monitoring wires and materials used to plug holes preventing radiation from escaping from underground caverns where nuclear weapons explode.

These small steps, delegates said, could lead to future agreements limiting nuclear tests, and perhaps paring superpower nuclear arsenals.

Mary Manning

India-USSR collaboration

New Delhi

NEGOTIATIONS have reached an advanced stage between India and the Soviet Union on the establishment of an international space centre in India. The idea was suggested by the Soviet Union in November 1986, and since then the Indian government has been examining the implications.

Science Minister Mr K.R. Narayanan recently told parliament that the Indian response to the Soviet offer will depend on "the need for international cooperation for peaceful uses of outer space consistent with our commitment to self-reliance". He said that discussions had taken place between the two countries and a decision would be taken "after all details regarding the scope and financial arrangements are defined, and the acceptability of such a centre by the international community is clearly established". Any arrangement between India and the Soviet Union on the space centre "will necessarily take into account the security implications and ensure that our own programme does not get affected".

The Soviet proposal is for an integrated space complex with launch facilities for manned and unmanned flights, laboratories for applied space activities, and training of scientists and astronauts from different countries. According to Indian officials, the details are still being worked out.

The Soviet Union is also awaiting the Indian response to its offer of a 1,000-MW nuclear power plant of the VVER type. India has accepted the offer in principle and negotiations are going on over reactor safety and international safeguard issues. According to the chairman of the Indian Atomic Energy Commission (AEC), Dr M.R. Srinivasan, import of the pressurized light-water reactor burning enriched uranium will not affect India's own nuclear programme based on CANDU reactors using natural uranium and heavy water. But critics, including former AEC chairman Dr Raja Ramanna, disagree. The Soviet Union, which recently leased one of its nuclear submarines to the Indian navy is, however, confident that the sale will go through.

K.S. Jayaraman

Correction: King Faisal prize

PROFESSOR M. Greaves, joint winner of the £50,000 King Faisal International Prize for medicine, is at the Leukaemia Research Fund Centre at the Institute for Cancer Research (see *Nature* 331, 293; 1988).

Japanese higher education looks to English universities

London

JAPANESE higher education could soon take a firm foothold in England if plans currently being discussed by the universities of Durham and Reading bear fruit.

Durham is hoping to lease a plot of land and buildings to Teikyo University, of Tokyo. The scheme would allow Japanese students to spend time in England, principally to learn English. The leased unit would be a part of Teikyo University and would have no formal links with Durham, although interaction between the two institutions would be actively encouraged. Teikyo originally approached Durham through the British Council. Durham was seen as particularly suitable because of its respected School of Oriental Studies and its location in the north-east of England, where the presence of a Nissan car-manufacturing plant has given rise to a sizeable Japanese community. A spokesman for the University of Durham stressed that the scheme is chiefly for the long-term benefits of both the Japanese and British students and staff.

The situation at Reading is very different. The university authorities are talking with an unnamed Japanese party that wants to establish a new university for children of European-based Japanese nationals. It is estimated that in Britain

alone there are some 22,000 Japanese nationals, including 6,000 children.

Since December 1985, Reading has been under pressure from the University Grants Committee to dispose of a ten-acre site deemed surplus to requirements, and isolated from the university's main 300-acre campus. The town-centre site contains 35 buildings, including the university's nineteenth century great hall and the microbiology department. The upkeep of the site is posing difficulties for the university, and last June it was offered for sale or lease. It was originally advertised in the United States, but the Japanese showed most interest. The present negotiations are the only ones being undertaken. A university spokesman estimated that up to 1,000 students could be taught on the campus, with the possibility of accommodation being shared with the Reading students. The Japanese university would be a separate entity.

Although financial pressure was the main motive behind the Reading move, the university does not accept criticisms that it is misusing the good name of the British university system for short-term financial expediency. If the two schemes prove successful, further lucrative deals between Japanese and British universities seem likely to follow. **Simon Hadlington**

New virus lands in United States

Washington

A SECOND human immunodeficiency virus (HIV) that causes AIDS (acquired immune deficiency syndrome) has turned up in North America. The new virus, HIV-2, was found in an AIDS patient living in New Jersey. Although epidemiologists have expected HIV-2 to show up in the United States, the Centers for Disease Control (CDC) in Atlanta do not believe that this first case signals a new epidemic.

HIV-2 was first found in West Africa by French researchers. The US patient, a West African who came to the United States in December 1987, was diagnosed as having cerebral toxoplasmosis, a secondary infection frequently found in AIDS patients. The patient had a negative enzyme immunoassay for antibodies to HIV-1. Because the patient came from West Africa, doctors at the University of Medicine and Dentistry of New Jersey (UMDNJ) where the patient was being treated decided to try an assay for HIV-2 being developed by Genetic Systems Corporation of Seattle, Washington. The HIV-2 immunoassay was positive. HIV-2 Western blot revealed the presence of core and envelope proteins. DNA ampli-

fication with HIV-1 and HIV-2 specific probes indicated that the infection was caused by HIV-2.

The patient had a history of sexual contacts while in Africa, and had received injections while hospitalized there. HIV-2 is assumed to be transmitted in the same way as HIV-1. Details of the case were reported in the CDC publication *Morbidity and Mortality Weekly Report* for 29 January (vol. 37, no. 3).

Stanley H. Weiss, formerly at the National Cancer Institute and now at UMDNJ, says this appears to be an isolated case. CDC, in collaboration with the Food and Drug Administration and other institutions, have now screened some 23,000 serum samples for the presence of HIV-2 antibodies. Of these, 35 (0.2 per cent) tested positive using Genetic Systems' enzyme immunoassay, but none of the 35 was confirmed using HIV-2 specific Western blot. All 35 positive samples came from a subgroup whose activities placed them at high risk for HIV-1 infection.

Weiss says the virus has been isolated from the infected patient, and is now being grown, so there will be sufficient quantities to sequence it. **Joseph Palca**

CEST chief named

London

THE man who will head Britain's latest effort to improve on its poor record of bringing science to the market-place was appointed last week. He is Dr Bob Whelan, the new chief executive of the Centre for



Dr Bob Whelan

Exploitation of Science and Technology (CEST), a body jointly funded by the government and a group of private companies. CEST's brief is to alert policy makers, industrialists and the research community to emerging areas of strategically important technologies (see *Nature* 330, 196; 1987).

Whelan, who is on indefinite secondment from his position as marketing director of PA Technology, part of a British-owned international management and technical consultant group, expects to recruit a staff of a dozen, including economists and market analysts as well as scientists. One of the first priorities of CEST, which is based at the Manchester science park, will be to attract a large number of smaller technology-based businesses into its membership.

Whelan expects that once his staff have pinpointed a strategically important area of science, companies and academic institutions with relevant expertise will be identified and a detailed plan of action drawn up. **Simon Hadlington**

Happy birthday, IUE

Paris

THE International Ultraviolet Explorer Satellite (IUE) celebrated its tenth birthday in geosynchronous orbit over the Atlantic Ocean on 26 January. IUE, which is operated by the European Space Agency, the British Science and Engineering Research Council and the US National Aeronautics and Space Administration, is used by astronomers all over the world to study ultraviolet radiation blocked out by the Earth's upper atmosphere. With a large number of important observations under its belt, including the pinpointing of supernova SN1987A, observation of the comet IRAS and the remote galaxy Fairall-9, IUE remains one of the most productive astronomical telescopes, despite its small (40 cm) diameter mirror. **Peter Coles**

Seventy-five Nobel laureates confer at Paris summit

Paris

AT the invitation of French president François Mitterrand, 75 Nobel laureates gathered in Paris last month for a conference entitled "Facing the 21st Century: Threats and Promises". The conference was the idea of Nobel peace laureate Elie Wiesel, who, paraphrasing Paul Valéry, explained that, "since the future is no longer what it used to be, it is our duty to approach it with caution".

The five themes, each tackled by a separate working group, covered peace and disarmament, human rights, the third world, science and technology and culture and society — themes whose globality sometimes accentuated the frequent association of Nobel awards with anglophone countries of the Northern Hemisphere. Only one Nobel laureate from a 'developing' country was present — physicist Abdus Salam from Pakistan, who now works in Trieste. Similarly, as Wiesel put it, the restriction of Nobel awards to a few disciplines meant that "biologists, doctors and novelists would have to talk about peace — but why not?"

Of the sixteen recommendations presented by Wiesel, two themes dominated: respect for human rights, including the right to education and technology; and the need to attack the polarization between the Northern and Southern Hemispheres perpetuated by the third world debt. The laureates called for an international summit specifically to address what Mitterrand called "the terrible observation" that it is "not the countries of the North which contribute to finance the South, but the reverse".

On disarmament, familiar paradoxes were aired. Although scientists had discovered the principles upon which destructive weapons have been built, responsibility for their deployment belongs to society as a whole. More than one speaker pointed out that whereas the Northern Hemisphere has known 40 years of peace since the bombing of Hiroshima and Nagasaki, continued 'conventional' wars have continued to disrupt the south.

The threat to world health posed by the AIDS (acquired immune deficiency syndrome) pandemic is an issue which many felt crystallize the problems of this *fin de siècle* and the central role of science in their solution. Joshua Lederberg, pioneer of genetic engineering, pointed out that "it is a pyrrhic victory for a virus to eradicate its host", but admitted that "each individual death of an infected person is a counter-example" to those who believe in natural mitigation. While genetic engineering could be "our most potent means of analysing viruses and developing

vaccines", he felt that technological improvements in transport, sanitation and vaccination have paradoxically left us more vulnerable to microbial attack.

It was John Vane, 1982 laureate in physiology or medicine, who brought the search for an AIDS vaccine into the arena of world economics. If a new AIDS therapy is developed, he claimed, it "will not be because pharmaceutical companies want to cure the disease. . . . but because they want to make a profit". Vane called for a truce in competition between companies searching for an AIDS treatment and for a change in legislation, to shift the onus of responsibility from manufacturer

to governments if a successful therapy produces toxic side-effects.

Current legislation would only prolong the delay between discovery and commercialization of an effective product, said Vane, who also challenged manufacturers to collaborate on research and to forgo profits. "Treatment with Retrovir currently costs \$200 per patient per week . . . and . . . 5 million Africans are already infected. Neither they nor their governments could afford to pay for costly treatment."

Aware that the conference could not hope substantially to change world politics, Wiesel felt that "the fact that the conference took place is in itself significant and important". Wiesel's Foundation for Humanity, set up in 1986 with his \$300,000 prize money, will organize a reunion in two years to study the problems highlighted in Paris.

Peter Coles

Employee accuses Stanford of health and safety violations

San Francisco

STANFORD University president Donald Kennedy has called for a university-wide review of health and safety practices, in response to a letter he received on 10 December from a high-ranking employee complaining of wide-ranging safety violations, mismanagement and cover-ups. Aware of the current sensitivity of any questions regarding research safety, Kennedy has arranged for the review to be carried out largely by people from outside Stanford, and has promised to make the results public.

The accusations come at a bad time for Stanford, with local distrust of university research running high, especially regarding safety issues. Community opposition to the expansion of research facilities has delayed the occupation of a research building at the University of California at San Francisco (see *Nature* 328, 626; 1987) as well as the construction of new research and animal buildings at Stanford (see *Nature* 329, 477; 1987).

Stanford has promised to publish its findings as an act of good faith towards the community — "a lot of short-term pain", said spokesman Bob Beyers, for a potential long-term gain in credibility.

Stanford officials refuse to name the employee who wrote the letter, but local press reports have identified him as an operations manager with the health and safety department, who resigned in December. Press reports have quoted him as saying he was ordered to keep quiet — Stanford officials say the investigation will deal also with these charges of cover-up and intimidation.

Kennedy has named a committee of five advisers to help evaluate the review, consisting of former US Secretary of Educa-

tion Shirley Hufstедler; Jere Goyan, dean of the school of pharmacy at the University of California at San Francisco and former commissioner of the Food and Drug Administration; Wolfgang Panofsky, former director of the Stanford Linear Accelerator Center; Stanford economics professor Kenneth Arrow, chairman of the Stanford faculty senate; and Victor Calvo, a former state assemblyman for the district including Stanford.

The first phase of the investigation will focus on specific allegations of safety violations in 10 campus buildings. None of the allegations has revealed immediate grave hazards, said Kennedy. Most involve storage of toxic chemicals and potential re-entry of toxic fumes vented through fume hoods. The 10-member auditing team from Stanford Research Institute's environmental health programme began to interview employees on 11 January, and has promised a preliminary report by the end of the month.

The second phase of the audit, for which a team has not yet been named, will investigate possible management improprieties, including cover-ups and employee intimidation. In a third phase, Kennedy plans to go beyond addressing the complaints in the employee's letter, and will appoint a campus-wide task force to evaluate the entire process by which safety policy is designed and implemented at Stanford searching for flaws that may lead to safety inadequacies. A special focus in this phase of the investigation will be the consideration given to safety when new buildings are being designed, an issue particularly timely as Stanford enters a phase of construction of new research buildings (see *Nature* 329, 281; 1987).

Marcia Barinaga

Cancer risk 'slightly higher' for UK nuclear test participants

London

A COMPREHENSIVE study of the health of men who took part in the UK atmospheric nuclear weapons tests in Australia and the Pacific between 1952 and 1968 has concluded that participation in the test programme has not had a detectable effect on life expectation nor on total risk of developing cancer. The results do however, suggest an increased risk of developing multiple myeloma and leukaemia.

The four-year study, published last week*, was undertaken at the request of the Ministry of Defence by a team from the National Radiological Protection Board and the Imperial Cancer Research Fund, led by Sir Richard Doll.

For the study, mortality and cancer incidence rates of 22,347 men who participated in the tests and experimental programmes were compared with a similar group of 22,326 servicemen and civilians not involved in the tests. In both groups, mortality was less than expected from national rates, and the rates were closely similar for cancer and for all causes combined (80 and 83 per cent respectively for participants and controls of that expected for cancer and 80 and 79 per cent respectively of that expected for all other causes).

The large difference between the two groups in the mortality from leukaemia and multiple myeloma (22 deaths from leukaemia and 6 from multiple myeloma in the participants against 6 from leukaemia and none from myeloma in the controls) was mainly due to the extraordinarily low rates for the diseases in the controls. Compared to national rates, mortality from the two cancers increased by 13 and 11 per cent respectively in the participants.

Examination of rates from cancer in different groups of participants, divided according to measured doses of external irradiation, failed to show any trend for increases in leukaemia, multiple myeloma or all neoplasms with increase in recorded dose.

The authors say: "The evidence related to multiple myeloma and leukaemia (other than chronic lymphatic leukaemia) is confusing, and on balance it is concluded that there may well have been small hazards of both diseases associated with participation in the programme, but their existence is certainly not proven, and further research is desirable."

The authors estimate that some 17 per cent of participants were not traced and were consequently excluded from the study. They say, however, that from subanalysis they are satisfied that any resulting bias is small.

Commenting on the study in last week's *British Medical Journal*, Martin Gardner, professor of medical statistics at the Medical Research Council's Environmental Epidemiology Unit, says: "The preferred conclusion so far must surely be that some leukaemias, and probably multiple myelomas, have resulted from radiation exposure during the tests. This is a stronger conclusion than the authors are prepared to reach because of the lack of certainty in the findings."

The Ministry of Defence, which commissioned the study after media publicity suggesting radiation-induced ill-health among test veterans, says it regards the results of the study "as fully vindicating its view that the radiological protection measures adopted were effective and that the chance of anyone suffering harm to health as a consequence of participation is extremely small". It has agreed to a recommendation from the study's authors to continue to observe mortality and cancer rates in the participants for a further ten years.

Simon Hadlington

* Darby, S.C. et al. *Mortality and Cancer Incidence in UK Participants in UK Atmospheric Nuclear Weapon Tests and Experimental Programmes* (National Radiological Protection Board — R214, HMSO, London).

New Indian nuclear watchdog will press for more openness

New Delhi

FOR the first time, Indian scientists have formed an independent non-governmental body to speak on nuclear issues of general concern and to keep a watch on the all-powerful Department of Atomic Energy (DAE).

The creation of the Indian Nuclear Society (INS), consisting of scientists, academics, intellectuals and representatives of the nuclear industry, ends the 40-year monopoly of DAE, on which Indians

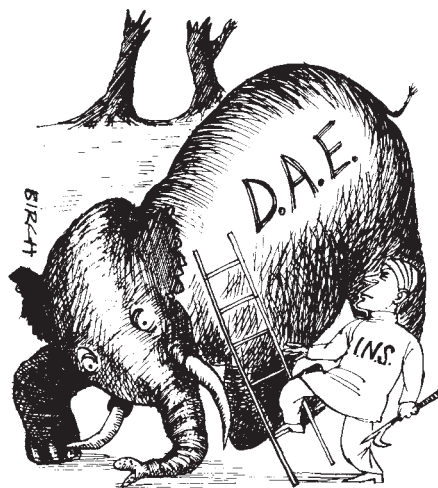
scientists, but to society as a whole."

Meaningful debates on issues of public concern are non-existent in India because of DAE's strict control over the flow of information from its projects, although all of them are supposedly for "peaceful purposes". DAE is also sensitive to criticism. A professor of the Jawaharlal Nehru University in New Delhi says the DAE caused his transfer from the Centre for Science Policy because of his critical articles.

Since its inception, DAE has been a closed establishment. It has been answerable directly to successive prime ministers and has operated under the provisions of the 25-year old Atomic Energy Act which, for example, prohibits any outside agency from even measuring the radioactive pollution near DAE's plants. In the past, DAE has been accused of either suppressing or providing incomplete details about leaks at the Tarapur plant, explosions in heavy-water plants, water pollution and death at a waste dump outside its nuclear fuel complex, as well as radiation hazards at its thorium plant in Kerala. Sites for nuclear stations or waste dumps are selected by DAE without any debate. In the absence of an independent body, DAE has a poor credibility rating. Scientists say that INS will in future provide a platform to discuss DAE's activities more openly.

According to Ramanaiah, the society will encourage nuclear energy and its applications and, at the same time, promote active participation of the public on nuclear issues. In carrying out its watchdog function, INS does not expect to embark on a collision course with DAE. Himself a firm believer in open policy, the new chairman of DAE, Dr M.R. Srinivasan, was instrumental in detaching the power wing from DAE and converting it into a public company last year.

K.S. Jayaraman



were forced to rely for any information — good or bad — connected with the atom.

The response to INS from the scientific community has been so great that 660 scientists and engineers, some of them ex-employees of DAE, enrolled as life members within a week of the formation of the society. Its chairman, Dr M.V.R. Ramanaiah, a retired scientist of the Bhabha Atomic Research Centre, says the society will function along the lines of the American Nuclear Society. "It will act as a watchdog and provide a forum for raising issues of relevance not only to nuclear

India since independence

SIR—In your leading article “Defining half a global problem” (*Nature* 327, 1; 1987) you say about poverty that “the only saving grace is that some among [the poor] (India in the past two decades, for example) occasionally break out of this depressing mould and find a way of managing their affairs constructively. If in the process of becoming semi-rich they think it necessary to cut down a few hundred hectares (or square kilometres) of primeval rain-forest, it will be found that local people are less regretful than Ms Bruntland.”

Like many of my countrymen, you have fallen for the rhetoric perpetuated by both the government and our so-called intelligentsia (of which scientists are a part) that India has made great strides since independence. A cursory glance at the achievements of China and India since independence reveals vast disparities in their achievements, whether *per capita* income, agriculture, science, education or health. Both countries started with similar problems at independence in the late 1940s. China, despite being cut off from the rest of the world and threatened by Japan, the United States and the Soviet Union, has been able to surpass India in almost all fields even though India has had free access to scientific development elsewhere in the world and has received large sums in foreign aid denied to its neighbour.

It would be more appropriate to compare the achievements of India with those of China rather than the rest of the developing world. The capitalist West, because of its morbid fear of communism, would like the developing countries to look up to India as the model for ‘freedom’ and ‘democracy’. Unfortunately ‘freedom’ has meant only freedom for a few to exploit and the rest to starve; for contractors to be permitted to decimate our forests; for business and industry to exploit the worker who produces goods consumed by only the upper two deciles. The process has meant inhuman living conditions in urban slums to which the neglected rural section, half of whom live below the poverty line, are forced to migrate. The best scientists and doctors, trained at public expense, emigrate to the West, or if they stay in India or return home, expect scarce resources to be diverted to the building of expensive laboratories and hospitals whose output in science and health not only leaves much to be desired but is often only a caricature of their Western counterparts whom they seek to emulate.

A market economy based almost entirely on human greed and devoid of state (let alone moral) controls has polarized the country’s society and its economy and sundered its fabric. Science and technology, instead of achieving its potential

for rejuvenating our society, has only pandered to the personal needs of those who control it and has been used more as a tool for exploitation than for development, under whose guise most sins are committed. One can only hope that China with its new-found ‘freedom’ does not follow suit.

N.H. ANTIA

Foundation for Medical Research,
84-A RG Thadani Marg,
Worli,
Bombay-400 018, India

Scientific truth

SIR—When one agrees with the premises and conclusions of an article it is tempting not to rock the boat about the reasoning used. There are very good reasons why twentieth century philosophy of science, under the malign influence of Popper through to Feyerabend, is profoundly hostile to science itself¹, as Theocharis and Psimopoulos (*Nature* 329, 595; 1987) correctly recognize. It is indeed unfortunate that many scientists, through ignorance, quote these philosophers approvingly. The most effective victories are those in which the losers unwittingly assist their opponents. But to suggest that the downgrading of science by philosophers is largely responsible for its loss of prestige (and consequently funding) is to overestimate the influence of philosophy. It is far more likely that stronger causes are the visible consequences of science’s misapplication: nuclear weapons, pollution and so on, together with the unfulfilment of the simplistic expectation that science ‘can solve all our problems’. Similarly, it is not the views of philosophers that are science’s worst enemy, but the proliferation of new age and fundamentalist cults, such as creationism, iridology and crystal therapy, that have no foundation in fact.

Theocharis and Psimopoulos, in common with their opponents, regret the inability to define the ‘scientific method’. But it does not matter: science attained its present level without such self-analysis, and there is no reason to suppose it cannot continue without it. Philosophy of science would become valuable to scientists only if it were able to provide a systematic replacement for the intuitive ‘leap in the dark’ involved in all discoveries. There are few less likely prospects than the formalization of intuition.

Nor is ‘scientific truth’ defined, even though it is defended against Kuhn’s view that competing theories come and go as arbitrarily as fashions². What Kuhn mystifyingly refuses to perceive is that the direction of change is not random, but is always towards better predictability: relativity predicts the outcome of dynamical events

more accurately than the newtonian view it succeeded, for example. This process of making ever better predictions is scientific progress, and it circumvents entirely the problem of defining scientific truth.

ANTHONY GARRETT

School of Physics,
University of Sydney,
Sydney, NSW 2006,
Australia

1. Stove, D.C. *Popper and After: Four Modern Irrationalists* (Pergamon, Oxford, 1982).
2. Kuhn, T.S. *The Structure of Scientific Revolutions*, 2nd Edn (University of Chicago Press, 1970).

SIR—Theocharis and Psimopoulos¹ miss the essential point when they state, under the sub-heading methodology, that the hapless student will have recourse to such (unreliable) things as random guess, arbitrary conjecture, subjective hunch, casual intuition, raw instinct, crude imagination and pure chance. Can this, they ask, be an adequate methodology by means of which to make new discoveries and beneficial applications? The history of most major break-throughs in scientific knowledge shows that it is.

Indeed, it is the element of surprise, the total negation of clear step-by-step logical thinking, that enables one’s patent attorney legally to establish an inventive step when claiming patent protection; a step clearly set down for example in the European Patent Convention¹ where it states in Article 56: “An invention shall be considered as involving an inventive step if, having regard to the state of the art, it is not obvious to a person skilled in the art.”

FRANK W. COUSINS

43 Emanuel House,
18 Rochester Row,
London SW1 1BS, UK

1. *Convention on the Grant of European Patents* 3rd edn (European Patent Office, Wila, Munich, 1985).

Colour blind

SIR—It seems that your reviewer of *Journal of Enzyme Inhibition* (*Nature* 329, 376; 1987) missed the point about colour illustrations that appear in this (and most other) Harwood and Gordon and Breach journals.

The colour plates are printed separately and inserted; a black and white reproduction is integrated into the text as a reference only (the legend points to the colour original). Sometimes the black and white plate, coming from a colour original, does not show appropriate contrast. But this is nothing to do with the paper quality, which will reproduce good contrast black and white original photographs well. Neither is it critical, as the full detail necessary appears in the colour plates.

JOHN GILLMAN

Harwood Academic Publishers GmbH,
PO Box 197,
London WC2E 9PX, UK

Whatever happened to the axion?

The continuing absence of the particle called the axion is an embarrassment for particle physicists, especially because it is improbable that its discovery need wait on new accelerators.

EVER since Dirac's daring suggestion, in 1929, that the symmetry of what seemed to him a properly relativistic wave equation for the electron made the existence of its positively charged partner, the positron, essential, particle physicists have grown used to the notion that new particles are more often predicted than discovered unexpectedly. Thus it was with Yukawa's mediator of the strong nuclear interaction (eventually identified as the pion), the anti-proton, the omega Ω^- meson, the charmed quark (otherwise the J/ψ) and, most recently and spectacularly, the heavy bosons that mediate the weak nuclear interaction known as Z^0 and $W^\pm$. The few surprises have been the recognition that the 'mesotrons' of cosmic-ray physics in the late 1930s included not merely the expected pions but the heavy analogues of electrons now called muons, the discovery of strange particles in the late 1940s and the later but less surprising discovery of the third generation of leptons, the tauon.

None of this is surprising. Finding a novel particle is not child's play, but almost always a matter of searching for a meagre signal against a rich background of distracting particle transformations. Rarely are the predictions of the existence of a new particle accompanied by estimates of the mass as accurate as those for the bosons of the weak interaction, with the consequence that people must go hunting through the energy spectrum before they find a particle that generally fits the bill. Only then can the properties of the particle be defined so as to tell the circumstances in which they are likely to be observed.

Unfortunately, this line of argument does not provide a suitable explanation for the continued absence of the particle called the axion, an electrically neutral object whose mass is probably the equivalent of at most a few electron-volts and which should be recognizable by its decay into two photons which, because the total energy available is limited, would be recognizable in the ultraviolet. On the face of things, detection should be a simple matter, not requiring the availability of great particle energy from a front-line accelerator. But the axion remains elusive.

That such a particle should exist stems from the observation, originally by Lee and Yang more than a quarter of a century ago, that parity (evenness or oddness of a

state) is not invariably conserved in transformations effected by the weak nuclear interaction. This entirely unexpected happening, mirrored a little later by the recognition that, in the decay of the 'strange' meson K^0 , the principle seems not to apply that the laws of physics are invariant when the direction (or algebraic sign) of time is reversed, seems however to be exclusive to the weak interaction. The strong nuclear interaction, by contrast, faithfully respects both parity and time reversal.

Particle physicists are entirely familiar with such states of affairs. If there is one rule for one set of circumstances and another for a second, those such as particle physicists whose goal is to find a unified description of phenomena speak of there being an underlying symmetry which has somehow been broken in the exceptional case. If, as seems sensible, they require that the results of calculations from the laws of physics should be independent of the arbitrary complex numbers the elementary equations may embody, which is what the notion of gauge symmetry entails, each broken symmetry requires the existence of a particle, or perhaps a family of particles. The axion emerges from such a chain of argument. Technically, it must be a boson, with integral, not half-integral spin, and it must have some mass, which may however be very small.

As will by now be familiar, axions are potentially convenient particles. Indeed, over the past few years, they have been widely canvassed as candidates for the WIMPs (weakly interacting massive particles) whose unremarkable presence in the Universe in very large numbers might provide the missing mass required to hold the Universe together. From time to time, there have been reports of ultraviolet background radiation from distant regions of the Universe suggestive of axion decay, but none has carried conviction. Curiously, it has fallen to astrophysicists rather than to cosmologists and particle physicists to pin down the properties of axions a little more precisely. Because of the weak interactions involved in thermonuclear fusion, stars should be prolific sources of axions, whence upper limits can be placed on their mass (unlikely to exceed 25 eV, or 0.001 per cent of the mass of an electron). But axions must certainly interact weakly with other matter, while the half-life of their spon-

taneous decay into a pair of photons must be long, probably exceeding the present lifetime of the Universe.

So what hope can there be of finding evidence of axions? Thomas W. Kephart and Thoman J. Weiler of the Vanderbilt University at Nashville, Tennessee, had an ingenious idea (*Phys. Rev. Lett.* **58**, 171; 1987) almost exactly a year ago. They suggested that because axions have mass, they should congregate together in clusters. Even though their lifetime is expected to be very long, there is no reason why their combined radiation of decay photons should not be perceptible, to say the least of it.

Indeed, Kephart and Weiler worked out in their formal calculation that, in spite of the very long lifetime of these objects, the sheer weight of numbers of axions in a cluster should give them a luminosity comparable with that of ordinary stellar matter. In short, people should have set out to look for diffuse patches of luminosity in the sky whose total light output might be comparable with that of the Sun, but which would seem much less intense because of their physical extent. (Because axions interact only weakly with other matter and only gravitationally with themselves, they are not easily cooled and thus condensed.) What the calculations showed is that axion light should be detectable from the Milky Way or from external galaxies such as Andromeda provided that axions account for one part in a million of the observable mass, which is modest by the cosmologists' expectations.

So what has happened in the intervening year? Nothing much, it seems. The great astronomy journals have not been filled with reports of axion searches, successful or otherwise. Nor, apparently, have people been much concerned to elaborate the calculations so as to suggest where in the sky searchers should most efficiently point their telescopes. One difficulty, that should eventually be a benefit, is that the radiation from an axion cluster might simply appear as a very narrow and very faint ultraviolet line that would escape detection so long as the mass is so ill-defined. But there is also a chance that the axion does not exist (there are other candidates for the role of the all-pervading WIMP), in which case there will be a scandal in particle physics. Should not that danger be taken more seriously?

John Maddox

Archaeology

Unearthing Byzantine politics

Richard Hodges

"POLITICAL history is rarely the province of an archaeologist", Professor R.M. Harrison asserts in his long-awaited account¹ of the excavations at Saraçhane, Istanbul. In this case, though, the site was selected for investigation precisely because it was at the interface of mainstream Byzantine history and archaeology. The results are of great importance, although not perhaps for the reasons uppermost in the excavator's mind at the start of the project.

In April 1960, grading operations at Saraçhane uncovered a large number of richly carved architectural blocks (see figure). Two carried an inscription, and these words were recognized as belonging to the 76-line epigram on the church of the martyr Polyeuktos.

This church was built by Anicia Juliana in AD 524–27. Anicia Juliana was the daughter of the Emperor Flavius Anicius Olybrius. She was a devout christian, disillusioned by politics of the Emperor Justin I and his celebrated successor Justinian, and set out to make an enduring statement about her times. The full epigram recorded in the *Anthologia Palatina* makes clear that the church dedicated to Polyeuktos, the Roman soldier who was martyred in AD 250, was built to eclipse the other buildings in the Byzantine capital. Anicia Juliana evidently erected the grand, domed church beside her palace, wherein the head of the martyr was kept.

The sixth-century Frankish chronicler, Gregory of Tours, was impressed by her, and tells how she tricked Justinian, who sought to deprive her of her riches, by casting all her gold into plaques which were affixed to the vaulted roof of her church. "My poverty is contained in this work", she told the great warrior-emperor, "do with it whatever you please". Justinian, it seems, was provoked to build an even greater monument — Hagia Sophia, which, unlike St Polyeuktos, has survived to the present.

Harrison's rescue excavations confirm the descriptions in these written sources. Evidently, St Polyeuktos was a grand building, lavishly decorated with marble panelling and brightly lit with multi-coloured windows. His excavations show that dumps were thrown against its walls in the eighth century and again in the tenth, and that three centuries later it was derelict. In 1204 the venetian crusaders pillaged some of its splendid ornaments, and after Constantinople fell to the Turks, it was destroyed and mosques were built within its once extensive precincts.

But the archaeologist and historian will be principally impressed by the incidental footnotes to this great history. Harrison is among the first to excavate a site of this kind thoroughly, bringing to light not only the sculptures but also objects as diverse as the pottery, lamp-chains, horse-shoes, bone objects and tiny bronze coins. These

artefacts tell a story of no less significance for the history of the Mediterranean.

In the past decade, several historians have drawn on the miscellaneous evidence for Dark Age Byzantium to show^{2,3} how the great fifth- to sixth-century ports and towns of Asia Minor dwindled to tiny, impoverished nuclei in the eighth and ninth centuries. Trade and commerce in the east Mediterranean seems to have declined dramatically, experiencing a nadir which endured until the later tenth century. This pattern contrasts to that in north-west Europe, where during this time a revival occurred almost as soon as direct Byzantine contacts fizzled out in the mid-seventh century⁴.

Yet the archaeology of Dark Age Byzantium is in its infancy. For instance, important excavations of wrecks off Turkey, including the seventh-century Yassi Ada ship⁵ and the eleventh-century ship at Serçe Liman⁶, leave no doubt that boat technology — our equivalent of space technology — did not falter in the Dark Ages. Nautical architects worked to improve ship design during these unknown years, while the ceramic amphorae, tablewares, glass drinking vessels and other objects discovered in the holds of these boats illustrate that simple craft technologies were also not lost. But who patronized these industries, and where?

The excavations at Saraçhane provide some answers to these questions. M.F. Hendy, in describing coins from Harrison's excavations, calls the assemblage of eighth- and ninth-century numismatic material "extraordinary" — coins exist in abundance in Constantinople whereas they are very rare in the rest of the Empire⁷. Similarly, objects as diverse as cooking-pots and lamp-chains associated with these coins illustrate that craftsmen were working for the palace-dwellers close to St Polyeuktos, and that the goods were plentiful enough to be discarded in some quantity.

Cyril Mango leaves no doubt that the city of Constantinople had a population of a million or more in the fifth century, yet experienced a staggering depopulation by the eighth century, when no more than tens of thousands lived there, surrounded by derelict ruins⁷. Even so, the city was still enormous by the standards of that time — Ephesus, for example, had

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

A marble pier-capital, with palm tree, from the Church of St Polyeuktos in Istanbul, first uncovered in April 1960. Photograph courtesy of Elizabeth Harrison.

1. Harrison, R.M. *Excavations at Saraçhane in Istanbul 1* (Princeton University Press, 1986).
2. Foss, C. *Am. J. Archaeol.* **81**, 469–86 (1977).
3. Hendy, M.F. *Studies in the Byzantine Monetary Economy, c.300–1450* (Cambridge University Press, 1985).
4. Hodges, R. & Whitehouse, D. *Mohammed, Charlemagne and the Origins of Europe* (Duckworth, London, 1983).
5. Bass, G.F. & van Doorninck, F.H., *Yassi Ada: A Seventh Century Byzantine Shipwreck* (Texas A&M University Press, College Station, 1982).
6. van Doorninck, F.H. *Int. J. naut. Archaeol.* **11**, 7–11 (1982).
7. Mango, C. *Byzantium* (Weidenfeld and Nicholson, London, 1980).
8. Foss, C. *Ephesus after Antiquity* (Cambridge University Press, 1979).

declined to little more than a fortified village⁸. In Constantinople, it seems the traditions and skills of Roman craftsmen were maintained and, despite the grave decline and dereliction, a vital connection linked classical and mediaeval civilizations. For these reasons, Harrison's excavations at Saraçhane permit archaeo-

logists to make assessments both of the political economies of the aristocracy in the age of Justinian, and of the obscure centuries which followed. □

Richard Hodges is in the Department of Archaeology and Prehistory, University of Sheffield, Sheffield S10 2TN, UK.

Meteorites

News from the early Solar System

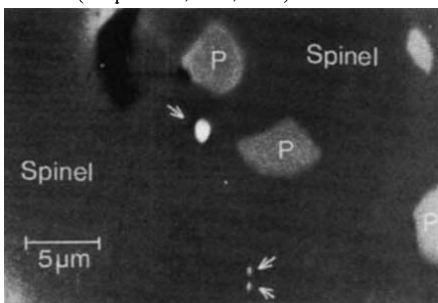
David Wark

AT the birth of the Solar System, some of the first solids to form in the solar nebula were refractory, Ca–Al-rich inclusions (CAIs)¹. These white, marble-sized CAIs were incorporated and preserved in carbonaceous chondrite meteorites which ultimately brought them to Earth. They contain large numbers of micrometre-sized metal nuggets (see figure) which are composed of the very refractory metals Mo, W, Re, Pt, Os, Ir and Ru in approximately the same relative proportions as occur in the Sun. Fegley and Palme have shown² that such alloys provide convincing evidence of high local temperatures of at least 1,300 °C in the solar nebula. Some CAIs also contain much larger opaque assemblages of OsRu grains, NiFe metal, magnetite (Fe₃O₄) and various other phases. In contrast to the nuggets, these larger refractory-metal assemblages do not reveal much about high-temperature events. But, as Blum *et al.* explain on page 405 of this issue³, they provide rich information about the lower temperature, post-formation history of the CAIs.

In a series of elegant laboratory experiments, Blum *et al.* demonstrate that the opaque assemblages (see their figures on pages 405 and 406) formed *in situ* in the CAIs by the oxidation of solidified globules of FeNi metal containing refractory metals in solid solution. Removal of Fe from the metal to form magnetite eventually caused grains of the less soluble metals Os and Ru to exsolve. Although this oxidation and exsolution model^{4,5} is not new, the work of Blum *et al.* places for the first time quantitative constraints on the temperature (600 °C) and oxygen partial pressure (pO_2 , six orders of magnitude higher than in solar gas) at which the assemblages formed.

The other achievement of the work is that it excludes, on the basis of Ru diffusion profiles in the NiFe metal that surrounds the exsolved OsRu grains, models⁶ in which pre-existing opaque assemblages were incorporated into still molten CAIs. The impetus for such capture models came from the fact that some opaque assemblages contain highly oxidized components, such as Fe₃O₄ and CaMoO₄, that are distinctly out of place among the

highly reduced, Ti³⁺-bearing silicates of the host CAIs. The opaque assemblages were therefore named *Fremdlinge*⁷ (strangers) by El Goresy, who suggested that they might contain information about extreme nebular environments, perhaps even about pre-solar environments in supernova ejecta. Although such capture models raised hopes that we might have stumbled upon some very exotic materials, the models were never quite able to account for the survival at 1,250 °C in the molten CAIs of the very volatile components (sulphides, Sn, Ge) that are not



Scanning electron micrograph showing a portion of a CAI containing typical refractory-metal nuggets (arrows). P, grains of perovskite (CaTiO₃) in spinel (MgAl₂O₄). The alloy nuggets contain (in weight per cent) ~24 Pt, 18 Mo, 16 Ru, 12 Os, 12 Ir, 11 Fe, 3 W, 2 Ni and 1 Re. Alloys like these, in which the refractory (Mo, W, Re) and noble (Pt-group) metals are combined, do not occur naturally on Earth.

uncommon in opaque assemblages.

We may not yet have heard the last of *Fremdlinge* however: there is evidence that there are at least some opaque objects foreign to their host CAIs, though not in the originally intended sense of having been incorporated ready-made. This evidence was uncovered by Bischoff and Palme⁸ who extracted and analysed several large intact opaque assemblages, and found to their surprise that the bodies possessed a completely different refractory-metal signature from that of the host CAI. They suggested that a few large, homogeneous but foreign metal grains were captured by the molten CAI and while cooling, underwent oxidation and exsolution.

The experiments of Blum *et al.* not only reproduce the textures observed in the natural opaque assemblages, but also are

able to place limits of 10 days to 10 years on the duration of the oxidation and exsolution process, depending on the particular temperature at which exsolution began. This is probably far too short a time (the cooling is too fast) for the process to have occurred in the well-insulated planetesimals from which meteorites originate, and implies instead that the CAIs were still drifting in the nebular gas. According to the experiments, the pO_2 of the gas surrounding the CAIs at this time was some six orders of magnitude higher than that of the approximately solar gas in which the host CAIs crystallized at about 1,250 °C. Because oxidation of the metal did not begin until the CAIs had cooled to about 600 °C, the pO_2 of the gas probably increased only gradually as the temperature fell. There is some independent evidence for such a gradual increase in pO_2 . Palme and I calculate⁹ that oxidized Fe²⁺ was introduced from the nebula into the rims of CAIs and olivine chondrules at a pO_2 roughly two orders of magnitude higher than solar gas at the intermediate temperature of about 900 °C.

The interesting question follows: if we are indeed seeing a trend of increasing pO_2 as the nebular gas cooled, what was the cause of the increase? Was oxygen being added, perhaps in the form of H₂O ice from the outer Solar System? Or, equivalently, was H₂ being preferentially removed, perhaps as part of the inevitable dispersal of the nebula as the Solar System evolved? It is conventional to speculate that intense solar wind outbursts during a 'T-Tauri' phase of the early Sun were the 'broom' which swept residual gas out of the solar nebula. Even assuming that to be true, however, it is still not known how far the formation of meteorites and accretion of planetesimals had progressed when the T-Tauri winds are supposed to have blown. Almost nothing is known for certain about the timing of the accretion of the inner (terrestrial) planets relative to that of the outer (gaseous) planets, or relative to the thermonuclear ignition of the proto-Sun, nor about the duration of each process. There is no shortage of problems, and investigations of meteorites continue to provide a touchstone for theoretical modelling of the origin and evolution of the Solar System, and of the star-forming process in general. □

1. Grossman, L. A. *Rev. Earth planet. Sci.* **8**, 559–608 (1980).
2. Fegley, B. & Palme, H. *Earth planet. Sci. Lett.* **72**, 311 (1985).
3. Blum, J.D., *et al.* *Nature* **331**, 405–409 (1987).
4. Palme, H. & Wlotzka, F. *Earth planet. Sci. Lett.* **33**, 45 (1976).
5. Wark, D.A. & Lovering, J.F. *Lunar Sci.* **IX**, 1214 (1978).
6. Armstrong, J.T., El Goresy, A. & Wasserburg, G.J. *Geochim. cosmochim. Acta* **49**, 1001–1022 (1985).
7. El Goresy, A., Nagel, K. & Ramdohr, P. *Proc. 9th Lunar planet. Sci. Conf.* 1279–1303 (1978).
8. Bischoff, A. & Palme, H. *Geochim. cosmochim. Acta* **51**, 2733–2748 (1987).
9. Palme, H. & Wark, D.A. *Lunar planet. Sci.* **XIX** (in the press).

David Wark is at the Max-Planck-Institut für Chemie, Saarstrasse 23, Postfach 3060, D-6500 Mainz, FRG.

GTP-binding proteins

Lithium affects G-protein receptor coupling

Alan H. Drummond

SINCE their discovery almost 40 years ago by Cade, lithium salts have been used effectively in the treatment of mania. It remains a mystery how a simple metal ion can exert such a profound psychopharmacological effect in the relative absence of systemic side-effects. The two theories that best explain the effect of this ion depend on the perturbation of the intracellular signalling molecules used by neurotransmitters: inositol lipid-derived mediators and cyclic AMP¹. On page 440 of this issue², Avissar *et al.* report that lithium, at therapeutic concentrations, can alter the function of a core feature of both of these signalling systems, namely GTP-binding proteins.

A substantial proportion of the body's different cellular receptors transmit their information into the cell by virtue of their ability to activate a relatively small, but growing, number of GTP-binding proteins³. Although various agents that can perturb G-protein function are known — for example, cholera and pertussis toxins and forskolin — little attention has been paid to the possibility that lithium ions might exert their therapeutically relevant effect at this site. Nevertheless, it has been evident for some time that the ion has the ability to affect G-protein-dependent phenomena such as receptor-activated inositol lipid metabolism and adenylate cyclase. Occupation of appropriate hormone and neurotransmitter receptors by their agonists leads to an exchange in the guanine nucleotide bound to the target G-protein such that GDP is replaced by GTP. Such changes can be monitored both by examining agonist-dependent changes in GTP binding and, more routinely, by monitoring the guanine nucleotide-dependent decrease in the affinity of the receptor for an agonist.

Inhibition

Avissar *et al.*², using both approaches, show that lithium can inhibit the coupling of both muscarinic cholinergic and β -adrenergic receptors to pertussis toxin- and cholera toxin-sensitive G-proteins, respectively, and that the concentration of lithium that exerts this effect (0.6 mM) is within the therapeutic range. Animals treated chronically with lithium show a similar effect that is reversible within 48 hours. The authors suggest additionally that this effect could explain recent data indicating an inhibition by lithium of both cyclic AMP formation and of the accumulation of inositol tetra-

kisphosphate (InsP₄), one of the many inositol phosphates that are formed following receptor-stimulated inositol lipid hydrolysis^{4,5}.

The most puzzling aspect of the ability of lithium to ameliorate the manic-depressive condition is its relatively selective action upon the central nervous system. In 1982, Berridge, Downes and Hanley proposed⁶ that lithium acts by interfering with neurotransmitter-stimulated inositol lipid metabolism. This hypothesis followed the pioneering work of Allison, Sherman and their collaborators (recently reviewed in ref. 1) who reported that inositol phosphate metabolism is blocked by low concentrations of the ion. The most persuasive element in this theory was that it explained why the brain is particularly sensitive to lithium: the supply of free inositol in the central nervous system is critically dependent on inositol phosphate catabolism (which is potently blocked by the ion) unlike the periphery, which has ready access to dietary inositol. Thus, the presumption was that after some time, and particularly in those regions of the brain that exhibit high activity (those that underlie the pathological condition?), lithium would trap much of the cellular inositol as inositol phosphates. The resulting inositol deficiency would, in turn, be reflected in a similar reduction in inositol-containing phospholipids; as a consequence, neurotransmitters that act through this signalling pathway would be unable to induce the formation of the inositol lipid-derived mediators, inositol 1,4,5-trisphosphate (InsP₃) and 1,2-diacylglycerol, that are essential for neurotransmission. The resultant inefficiency in neuronal communication would be particularly localized to areas of the brain that were highly active before drug treatment. If, as seems likely, the manic state results from this hyperactivity, an improvement in the clinical condition might then be expected to occur.

The data of Avissar *et al.* are important in that they suggest another potential biochemical end-point for the action of lithium that is affected by therapeutically relevant concentrations. They are preliminary, however, and further work is

Corrigendum

In the obituary of John H. Northrop (*Nature* 329, 396; 1987), the year that the Princeton branch of the Rockefeller Institute closed, and that Northrop moved to Berkeley, was 1949–50, not 1938–39 as stated. □

needed to establish whether this interaction can yield a preferential effect on the brain analogous to the clinical picture: all cells in the body exhibit G-protein-dependent phenomena and it is not immediately evident why, if the effect is present in lithium-treated manic patients, the function of peripheral tissues should not be affected.

In recent years, the ability of lithium to amplify inositol lipid-linked signalling by inhibiting inositol phosphate catabolism has been of enormous benefit in the detection of responses in tissues and cells that contain only a few receptors⁶. Direct effects of lithium on the levels of the important second-messenger molecules derived from the pathway are, however, only rarely seen⁷ and, in these examples, the physiological consequence was unclear. The work of Irvine and his collaborators^{8,9}, some of which was reported in a recent issue of *Nature*, suggests that InsP₄ may be involved in the regulation of intracellular calcium homeostasis by facilitating influx of the cation (see the News and Views article¹⁰ by Jennifer Altman).

Unique action

Moreover, Batty and Nahorski⁵ have recently demonstrated that the accumulation of InsP₄ is decreased by lithium in brain slices that have been stimulated by cholinergic agonists. Studies in peripheral tissues have, thus far, failed to reproduce this effect and, although the mechanism underlying it remains to be established, it is possible that the unique action of lithium on the central nervous system might be related to its ability to decrease InsP₄-dependent effects on calcium influx.

Attempts to extrapolate these lithium-sensitive biochemical events to the action of this ion in manic patients will be viewed, justifiably, with some caution by clinicians. It may be, in fact, that the most elegant proof that a relationship exists between the two phenomena will emerge if some of the current pharmaceutical industry interest in drugs targeted against inositol phosphate catabolism results in another active anti-manic drug. □

1. Sherman, W.R. in *Inositol Lipids in Cell Signalling* (eds Michell, R.H., Drummond, A.H. & Downes, C.P.) (Academic, London, in the press).
2. Avissar, S., Schreiber, G., Danon, A. & Belmaker, R.H. *Nature* 331, 440–442 (1988).
3. Gilman, A.G. *Rev. Biochem.* 56, 615–649 (1987).
4. Newman, M.E. & Belmaker, R.H. *Neuropharmacology* 26, 211–217 (1987).
5. Batty, I. & Nahorski, S.R. *Biochem. J.* 247, 797–800 (1987).
6. Berridge, M.J., Downes, C.P. & Hanley, M.R. *Biochem. J.* 206, 587–595 (1982).
7. Drummond, A.H. & Raeburn, C.A. *Biochem. J.* 224, 129–136 (1984).
8. Irvine, R.F. & Moor, R.M. *Biochem. J.* 240, 917–920 (1986).
9. Morris, A.P., Gallacher, D.V., Irvine, R.F. & Petersen, O.H. *Nature* 330, 653–655 (1987).
10. Altman, J. *Nature* 331, 119–120 (1988).

Alan H. Drummond is at British Bio-technology Ltd, Watlington Road, Cowley, Oxford OX4 5LY, UK.

Palaeontology

Early evolution of birds

Joel Cracraft

OUR understanding of the evolutionary history of the main taxonomic groups relies heavily on knowledge provided by fossil organisms. Reciprocally, it is usually not possible to obtain an accurate assessment of the phylogenetic relationships of fossil taxa without a prior hypothesis based on extant organisms. Fossil organisms can provide information on character-state distributions not found among living taxa, but without diagnostic characters, fossil specimens can be little more than old scraps of bone. Thus, palaeontology's great contribution to the study of the history of life is not so much that it adds a time dimension, although that contribution is undeniably important, rather that it yields copious character-state data that can be used to reconstruct genealogical relationships. When new, critical fossils are discovered, palaeontologists are justified in kicking up their heels. One cause for celebration is a new fossil bird from the Lower Cretaceous of Spain (Las Hoyas), which is the subject of a report by J. L. Sanz, J. F. Bonaparte and A. Lucas on page 433 of this issue¹. Although this find does not lead us to restructure our ideas on early avian relationships, it clarifies our knowledge of character evolution and provides important new interpretations regarding the early diversification of birds (Fig. 1).

The Late Jurassic species *Archaeopteryx lithographica* is widely acknowledged to be among the most important fossils in demonstrating a phylogenetic link

between two major groups of organisms. Indeed, all discussions of the origin of birds begin with the data provided by *Archaeopteryx*²⁻⁴. Primarily because it possesses feathers, the avian relationships of *Archaeopteryx* have seldom been questioned, but its more distant affinities to one or more reptilian groups have been

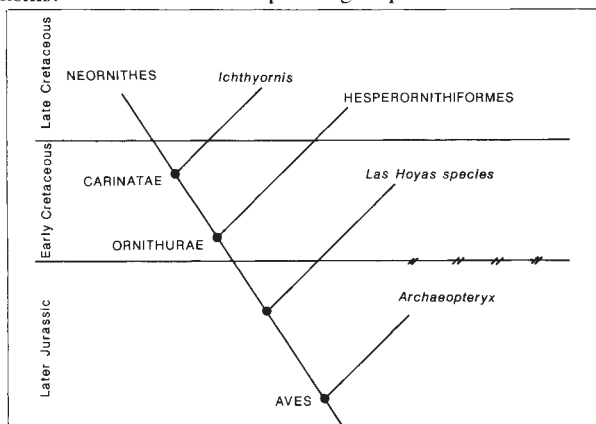


Fig. 1 Phylogenetic hypothesis for some of the taxa relevant to the early history of birds. Present information on the Las Hoyas species suggests it is the sister-group of the Ornithurae.

strongly debated⁵. For the past decade or so, an increasing amount of data point to the conclusion that *Archaeopteryx* and all other birds are more closely related to some theropod dinosaurs than to any other reptilian taxon⁶⁻⁸. This important finding provides a critical comparative foundation for interpreting evolutionary character transformations that took place during the early history of birds and has led to my phylogenetic hypothesis for some of the main avian lineages⁹.

The Las Hoyas specimens can be

interpreted relative to this scheme of relationships. As Sanz and colleagues note¹, the new fossil bird is seemingly not so primitive as *Archaeopteryx* nor as advanced as all other birds (collectively called the Ornithurae). The Las Hoyas fossil and the Ornithurae share a strut-like coracoid and pygostyle, both absent in *Archaeopteryx*. The Ornithurae have fused pelvic elements, complete fusion of the metatarsals, and complete fusion of the tarsals distal to the metatarsals. The Las Hoyas specimen, in contrast, lacks these features, as does *Archaeopteryx*. Given this character distribution, the conclusion of Sanz *et al.*, that the Las Hoyas specimen is the sister-group of the Ornithurae, appears justified (see Fig. 2).

If this phylogenetic hypothesis is correct then the Spanish discovery implies several anomalous character transformations and necessitates refinements in our interpretation of the evolution of the Late Cretaceous toothed divers, the hesperornithiforms. The apparent anomalies in the Las Hoyas specimen are that it is said to possess first, an astragalus and calcaneum that are not fused to the tibia and second, an unfused metatarsus. Remarkably, both of these features seem to be more primitive than the condition in *Archaeopteryx*⁵, and the latter character is paralleled to some extent in the Late Cretaceous enantiornithine¹⁰ birds (see also ref. 9). Whether these differences will be resolved following a more detailed description of the Las Hoyas material remains to be seen. In any case, to accommodate these character conflicts it may be necessary to postulate that fusion, or possibly even loss of fusion, of the tarsal and metatarsal elements occurred more than once.

The order Hesperornithiformes includes

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

a few flightless species which are highly specialized for diving¹¹. Despite possessing many highly modified features associated with the loss of flight and locomotion through an aquatic medium, hesperornithiforms have many primitive characters that exclude them from the carinates^{9,12}. Although it has been commonly assumed that hesperornithiforms are secondarily flightless, my cladistic analysis⁹ of available data does not support this hypothesis, chiefly because there are no more primitive taxa with characters indicating an ability to fly. The Spanish specimen could resolve this issue.

If the relationships of the Las Hoyas bird are as envisioned by Sanz *et al.*, then the hesperornithiforms were probably secondarily flightless. A pygostyle, a strap-like coracoid having a well-defined articulation with the sternum, and a furcula with a hypocleidium found in the Spanish material all suggest that the bird could fly. At the same time as the Las Hoyas specimens clarify the evolutionary history of the hesperornithiforms, they also require avian palaeontologists to

re-evaluate the diagnostic characters of the carinates themselves. Many of these characters will undoubtedly be found to define a larger group, namely all birds other than *Archaeopteryx*. A further implication of this result is that many morphological features characterizing modern birds arose very early in the history of birds. □

1. Sanz, J.L., Bonaparte, J.F. & Lacasa, A. *Nature* **331**, 433–435 (1988).
2. Beer, G.R. de, *Archaeopteryx lithographica; a Study Based on the British Museum Specimen* (Br. Mus. Nat. Hist., London, 1954).
3. Hecht, M.K., Ostrom, J.H., Viohl, G. & Wellnhofer, P. *The Beginnings of Birds* (Freunde des Juras-Museums, Eichstätt, 1985).
4. Heilmann, G. *The Origin of Birds* (Witherby, London, 1926).
5. Ostrom, J.H. *Biol. J. Linn. Soc.* **8**, 91–182 (1976).
6. Padian, K. *Proc. N. Am. Paleont. Conv.* **2**, 387–392 (1982).
7. Thulborn, R.A. *Zool. J. Linn. Soc.* **82**, 119–158 (1984).
8. Gauthier, J. *Mem. Calif. Acad. Sci.* **8**, 1–55 (1986).
9. Cracraft, J. *Paleobiology* **12**, 383–399 (1986).
10. Walker, C.A. *Nature* **292**, 51–53 (1981).
11. Martin, L.D. & Tate, J. Jr *Smithsonian Contr. Paleobiol.* **27**, 35–66 (1976).
12. Martin, L.D. in *Perspectives in Ornithology* (eds Brush, A. & Clark, G.A. Jr) 291–338 (Cambridge University Press, New York, 1983).

Joel Cracraft is in the Department of Anatomy and Cell Biology at the University of Illinois, Chicago, Illinois 60680, USA.

Biological modelling

What's evolving in artificial life

Doyne Farmer and Stuart Kauffman

IF we ever discover extraterrestrial life, the implications for biology will be enormous. We would be forced to confront the unavoidable parochialism of a science whose body of knowledge is built on only one variety of life, namely, life on Earth.

Contemplating the possibility of extraterrestrial life raises some basic questions: which features of life are universal? Which are specific to Earth — how much of the life that we see around us is a 'frozen accident' of evolution? This immediately leads to a host of more specific questions: must life be based on carbon chemistry; must the genotype–phenotype distinction

exist; do life forms need to be organized into species, with branching phylogenies?

We may or may not encounter extraterrestrial life, but whether or not we do, many of the same questions might be addressed by studying 'artificial life'. With advances in software, hardware and 'wetware', it may be possible to create lifelike 'organisms'. They might be carbon-based, created inside a laboratory, or robots, built out of silicon-based chips, or abstractions that exist only inside a computer. Artificial organisms can provide a laboratory to study various fundamental questions about life that cannot easily be

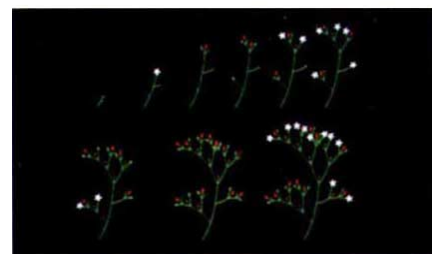
A simple recursive genetic description consists of a set of rules that define the pattern of change across generations. The following is a simple example:

- (1) A → CB
- (2) B → A
- (3) C → DA
- (4) D → C

Once the initial state is determined, each successive state is determined accordingly.

Time	Structure	Rules applied (L to R)
0	A	(initial 'seed')
1	C B	(rule 1 replaces A with CB)
2	D A A	(rule 3 replaces C with DA and rule 2 replaces B with A)
3	C C B C B	(rule 4 replaces D with C and rule 1 replaces the two
4	 (etc)	As with CBs)

More complex rules can include branching, and can lead to the type of recursive patterns represented in the figure.



Artificial beauty — computer-generated life forms can demonstrate patterns of morphogenesis. Below, development of a sympodial inflorescence (*Lychnis coronaria*) generated by a DOL-system (pictures courtesy of Chris Langton, Los Alamos National Laboratory).

addressed in ordinary biological experiments. Included in this are attempts to analyse the fundamental features of the 'logic of life', and to assess the conditions under which complex systems made of simpler components can self-organize and adapt in complex environments.

An interdisciplinary group met recently* for the first conference on artificial life. Models of the origin of life, of morphogenesis and of co-evolution in simulated environments were discussed. The first set of models tried to account for the spontaneous emergence of self-reproducing systems in simulations of the prebiotic environment.

The second set of models followed a line of thought that dates back to Goethe. Richard Dawkins (University of Oxford) and others presented elegant examples of morphogenesis of animal and plant-like forms, by the recursive application of simple developmental mechanisms. Slight modifications in the parameters can lead to slight alterations in the resulting morphologies in some cases, but large alterations in others. Thus, evolution explores a space of parameter values which generate a family of recursively related forms. However, it remains unclear how to incorporate a natural notion of fitness into these models.

Artificial ecosystems presented at the conference represent a third approach to the study of artificial life forms. Various models were presented: abstract animals living on a two-dimensional grid, with selection occurring on high-level properties such as 'aggression'; a model for the social interaction of bees; a 'movable finite automata' world in which the laws of 'chemistry' cause a virus molecule to self-

*21–25 September 1987 held at the Los Alamos National Laboratories, New Mexico.

assemble and then penetrate a cell wall; co-evolving 'worms' which compete for food in a world determined by the diffusion equation and the interactions between worms; and 'core wars', a computer environment in which computer programs compete with each other to dominate the core memory of their host machine. A virtue of many of these models is that the notion of 'fitness' is only implicit in the rules, in contrast to the usual models in which fitness is externally imposed. As Howard Pattee said, "Organisms do not evolve in an environment; rather they evolve their environment". These artificial environments may someday provide the laboratories for the study of 'natural' selection.

Although von Neumann demonstrated that it is possible to build interesting self-replicating structures in an abstract environment, his model and the others that have followed rely on a sentient creator, with no possibility of spontaneous emer-

gence, and little or no stability to perturbations. The most important problem for research on artificial life is how to design worlds of self-reproducing interacting entities with emergent collective properties, that are robust to noise, and have the spontaneous capability to co-adapt and produce functional structures.

The situation in the field of artificial life might be likened to that of thermodynamics in the early part of the nineteenth century. We have vague intuitive ideas what the important elements are, but no precise definitions or theories. The experimental or theoretical framework is still not well formulated, but at this conference there was a palpable sense that we are on the trail. We may see real progress in this area within the next decade. □

Doyle Farmer is at Los Alamos National Laboratory, Los Alamos, New Mexico 87545; Stuart Kauffman is in the Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia 19104, USA.

Pharmacogenetics

Enigmatic variations

Jeff Idle

"No-one supposes", wrote Darwin¹, "that all the individuals of the same species are cast in the same actual mould. These individual differences are of the highest importance to us, for they are often inherited." The genetic complexities of one of the commonest ways in which individuals differ from each other are unravelled by Gonzalez *et al.*, a group of US and Swiss authors, in a paper on page 442 of this issue². These studies demonstrate the genetic origins of debrisoquine metabolism deficiency, which affects 5–10 per cent of people.

People who suffer from this deficiency often experience profound pharmacological or toxic responses when prescribed ordinary doses of a growing number of drugs³. This debrisoquine 4-hydroxylase polymorphism is of importance not only to the affected individual but also to the pharmaceutical industry and drug-regulatory authorities. If, for example, a new drug, after millions of pounds have been spent on its development, turns out to be metabolized primarily by this cytochrome P450 enzyme, P450db1, then its clinical use and thus its commercial future will be in jeopardy⁴. It would be hard to establish significant clinical use of a new drug whose plasma profile, and thus efficacy and safety, were capricious. Darwin was not to witness the echoes of his own perspicacity: individual differences are of the highest importance to us not only because through them we perceive better the workings of heredity and evolution, but also because, variety being a natural consequence of

evolution, it has been difficult to produce a drug appropriate to everyone. The metabolic oxidation of drugs and other alien chemicals can be highly variable from person to person.

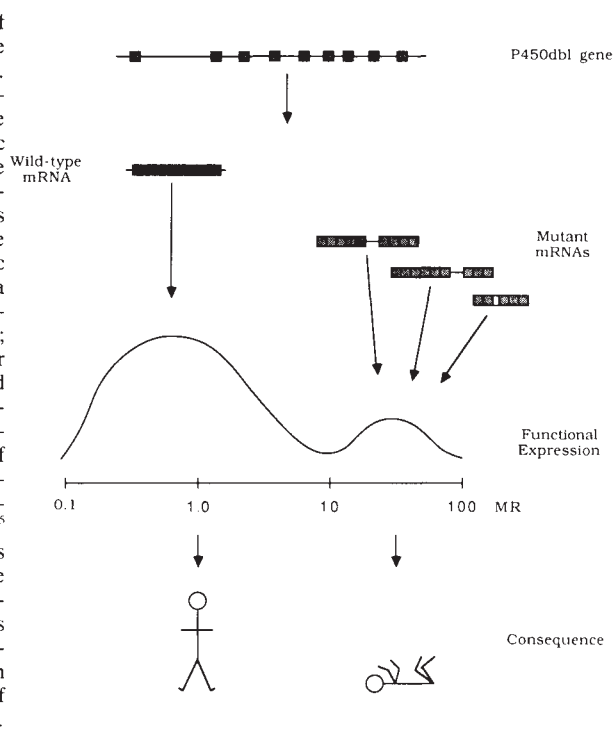
Elucidation of the mechanism of the debrisoquine deficiency, the first P450 polymorphism described, is rooted in serendipity. A professor of pharmacology,

who, together with several colleagues, took a small dose of the antihypertensive drug debrisoquine as part of a metabolic study, suffered an excessive fall in his blood pressure. Urine analysis, together with a wider population study, revealed⁵ that he and two others had an inherited defect in debrisoquine 4-hydroxylation and were the archetypes of the now well-established 'poor metabolizer' phenotype.

A decade of effort in many laboratories has established the practical and clinical importance of this common human phenotype. Hitherto it was thought likely that poor metabolizers originate from a single mutation. What Gonzalez *et al.* now show² is the presence of at least four mutations which all give rise to the poor metabolizer phenotype. Moreover, somewhat surprisingly, these mutations occur not within the coding exons of the gene, but in the introns, giving rise to a series of incorrectly spliced messenger RNAs that are unable to yield immunodetectable protein. As the authors point out, these defects are unusual, difficult to detect and will require sequencing of the complete normal human P450db1 gene, including introns, exons and flanking regions, before the source of aberrant splicing of the pre-messenger RNA in poor metabolizers can be clarified.

Sadly, this may delay the development of new strategies for unequivocal phenotyping and detection of poor metabolizers using techniques that do not involve drug administration. A single mutation would have led more briskly to the development of a linked restriction-fragment length polymorphism analysis of DNA derived from a small blood sample. This notwith-

Molecular characterization of the human P450db1 gene according to Gonzalez *et al.*². Normal splicing of pre-messenger (m)RNA gives the wild-type mRNA. Genetic mutations, however, cause splicing defects and thus production of mutant mRNAs that yield no detectable P450db1 protein. This genetic polymorphism is detected as a bimodal population distribution of metabolic ratio (MR; the ratio in urine between per cent dose as debrisoquine and per cent dose as 4-hydroxydebrisoquine). Poor metabolizers, arising from one of several mutations and consequent aberrant splicing, produce negligible metabolite⁵ and thus commonly have MRs between 20 and 100. These people are at high risk of certain adverse drug reactions and at low risk of some chemically induced diseases which involve the generation of reactive oxidative metabolites.



standing, it is rapidly emerging that the biology of the cytochrome P450 family is not simple, hardly surprising for a superfamily that comprises eight known gene families, each consisting of 2–20 discrete cytochromes⁶. Not only have we evolved a very complex defence against the constellation of 'chemical predators' in the environment, but it is also likely that several essential pathways of endogenous metabolism, and therefore many diseases, will involve some members of the cytochrome P450 family, probably the most highly conserved and invariable ones. In contrast, P450db1, whose gene Gonzalez *et al.* show can exist as one of several stable mutants, is probably concerned with the metabolism of environmental and dietary chemicals.

This contention is supported by the epidemiological associations between debrisoquine hydroxylation phenotypes and some chemically induced diseases including lung cancer⁷, bladder cancer⁸, hepatocellular carcinoma⁹ and endemic Balkan nephropathy¹⁰. Clearly, this P450 polymorphism is a locus for variable interaction between host and environment which, by selecting for one individual rather than another (probably in our earlier mammalian predecessors), has helped to fuel evolution. The molecular characterization of the polymorphism by Gonzalez *et al.*² will permit a much closer inspection of the relationships between environmental chemicals and genetic susceptibility to disease.

One final point of interest tucked away in the paper² is the use of monkey kidney fibroblast (COS) cells as vectors for expressing cloned P450 complementary DNAs. In essence, cells that have no drug-metabolizing capacity become, with the aid of viral genes, factories for the production of a single P450 isozyme. Armed with such cells it is easy to envisage how new drug candidates or chemical carcinogens could be screened for polymorphic metabolism. In simple experiments it would be possible to predict which drugs and chemicals are inappropriately metabolized by 5–10 per cent of the population, and therefore help to pre-empt the clinical, social and economic consequences of the deficiency. □

1. Darwin, C.R. *On the Origin of Species by Means of Natural Selection* (Murray, London, 1859).
2. Gonzalez, F.J. *et al.* *Nature* **331**, 442–446 (1988).
3. Idle, J.R. & Smith, R.L. *Proc. 2nd Wld. Conf. Clin. Pharmac. Ther.* **148** (1984).
4. Idle, J.R. *et al.* *Lancet* **i**, 1388 (1983).
5. Mahgoub, A. *et al.* *Lancet* **ii**, 584 (1977).
6. Nebert, D.A. & Gonzalez, F.J. *A. Rev. Biochem.* **56**, 945 (1987).
7. Ayes, R. *et al.* *Nature* **311**, 169 (1984).
8. Kaisary, A. *et al.* *Cancer Res.* **47**, 5488 (1987).
9. Idle, J.R. *et al.* *Cancer Lett.* **11**, 331 (1981).
10. Ritchie, J.C. *et al.* in *Current Research in Endemic (Balkan) Nephropathy* (eds Strahinjic, S. & Stefanovic, V.) 23–27 (Nis University Press, 1983).

Jeff Idle is Reader in Pharmacogenetics and Wellcome Trust Senior Lecturer, Department of Pharmacology and Toxicology, St Mary's Hospital Medical School, London W2 1PG, UK.

Particle physics

Observation of double β -decay

Michael S. Witherell

THE first direct evidence of the double β -decay of a nucleus is reported by S.R. Elliott, A.A. Hahn and M.K. Moe in a recent issue of *Physical Review Letters*¹. Double β -decay is an extremely rare form of nuclear decay in which two electrons are emitted, rather than one as in normal β -decay. Elliott *et al.* observed the decay of the isotope ^{82}Se into ^{82}Kr plus two electrons and two antineutrinos. The measured half life of $1.1^{+0.8}_{-0.3} \times 10^{20}$ yr makes this the rarest natural decay process yet

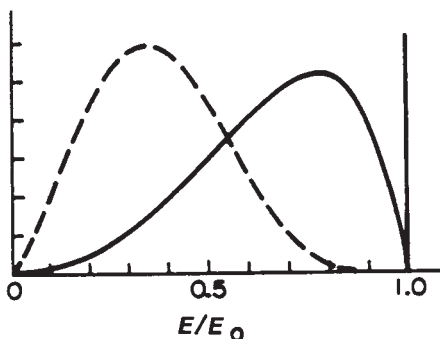


Fig. 1 Ideal two-electron energy (E) spectra for nuclear double β -decay with two neutrinos (broken line), with a majoron (solid line) and neutrinoless (vertical ' δ -function'). E_0 is the total decay energy (see text for details).

observed directly in the laboratory.

The ultimate goal of experiments studying double β -decay is to search for another process, neutrinoless double β -decay $\beta\beta(0\nu)$, which would mean a fundamental change in our picture of the neutrino. The neutrino was first observed indirectly as the particle that carries away the unseen energy in normal β -decay. In $\beta\beta(0\nu)$, if it exists, a nucleus emits two electrons with no antineutrinos. Such a decay would constitute a violation of lepton-number conservation, one of the few conservation laws in particle physics thought to be absolute. Lepton number is defined to be +1 for electrons and neutrinos, –1 for their antiparticles, positrons and antineutrinos. In the standard model of particle physics, any decay which produces an electron must also produce an associated positron or antineutrino. The neutrino in this model is said to be a Dirac neutrino.

In 1939, W.H. Furry showed² that $\beta\beta(0\nu)$ would be the dominant double β -decay process if the neutrinos were 'Majorana' particles rather than Dirac particles. In the Majorana theory of neutrinos, the neutrino and the antineutrino are the same particle. The only distinction is that neutrinos are left-handed, that is, their spin points opposite to their velocity, and antineutrinos are

right-handed. If neutrinos are massless particles, there is no way to distinguish the Dirac neutrino, with separate particle and antiparticle, from the Majorana neutrino. Thus, the observation of $\beta\beta(0\nu)$ would not only demonstrate the violation of lepton number conservation, but would also prove that the neutrino's mass is not exactly zero.

To date, there are no experiments in which neutrinoless double β -decay has been observed. The most sensitive searches look for the isotope ^{76}Ge decaying into ^{76}Se plus two electrons. This isotope is an 8 per cent component of natural germanium, an element commonly used to make detectors for nuclear physics. As shown in Fig. 1, the signature of $\beta\beta(0\nu)$ would be the observation of events in which the two electrons are emitted with a total energy of exactly 2.04 MeV. The fortuitous coincidence that germanium detectors with very good energy resolution contain an isotope that is a double β -decay candidate makes it possible to search for this decay with unprecedented sensitivity. At present, experiments show that the half life for this decay must be greater than 6×10^{23} yr. Using the existing calculations for the nuclear physics, this result says that if the neutrino is a Majorana particle, its mass must be less than 1 eV. For comparison, the particle with the lightest non-zero mass is the electron, with mass of 5.1×10^5 eV.

The main reason for the interest in double β -decay with neutrinos $\beta\beta(2\nu)$, the process observed by Elliott *et al.*, is to check the nuclear physics involved in interpreting the results for $\beta\beta(0\nu)$. Because the $\beta\beta(2\nu)$ process is allowed within the standard model of particle physics, its rate is calculable, in principle. The difficulty is that the transition rate depends on the complicated nuclear physics of the parent and daughter nuclei. In particular, calculations^{4–6} of the

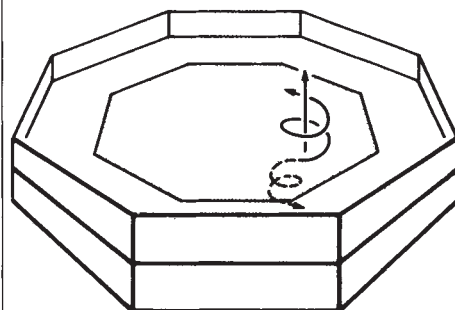


Fig. 2 Idealized neutrinoless double β -decay event in the time projection chamber of Elliott *et al.* showing two spiralling tracks originating from a single point. (Courtesy of A.A. Hahn).

half life for $\beta\beta(2\nu)$ decay in ^{82}Se are of the order of 10^{19} yr. The first indications that the half life is actually much longer came from geochemical experiments^{7,8} that measured half lives of around 1.3×10^{20} yr. In these experiments, the experimenters look for the daughter nucleus ^{82}Kr , which is a noble gas, occluded in rocks of known age containing ^{82}Se .

The experiment of Elliott *et al.* confirms the geochemical result using a time projection chamber (TPC) to observe individual double β -decays. The TPC is an ingenious device developed by D. Nygren of the Lawrence Berkeley Laboratory that has been used in many experiments to detect particles ranging in energy from 1 to 20,000 MeV. It consists of a large gas volume contained in combined electric and magnetic fields. A high-energy particle produces many ionization electrons along its trajectory, which is a spiral because of the magnetic field. These ionization electrons drift to the end of the chamber, where they are detected and their time and position accurately recorded. This position and the time it took to drift to the end are used to measure the exact position at which each ionization electron was produced. The full trajectory of the high-energy particle can be reconstructed from these measurements, and the curvature in the magnetic field is used to measure the momentum. For the particular use in the double β -decay experiment, the requirement that two electrons start at exactly the same point (Fig. 2) is used to remove many common background processes that can fake $\beta\beta(2\nu)$.

As shown in Fig. 1, the energy spectrum for $\beta\beta(2\nu)$ is not a narrow peak, but a broad continuum. The ^{76}Ge experiments, with precise energy resolution, are ideally designed for the neutrinoless decay peak. The TPC is more sensitive for the two-neutrino decay, however, because it observes each electron separately. The energy spectrum reported by Elliott *et al.* agrees well with the continuum expected for double β -decay, and is quite different from the dominant background. The measured half life of $1.1^{+0.8}_{-0.3} \times 10^{20}$ yr confirms the geochemical results, and establishes that the decay is an order of magnitude rarer than calculations predicted. Nuclear theorists must now reevaluate their calculations to understand this large discrepancy. It is important to see what effect the change will have on the Majorana neutrino mass. If the calculated rate for $\beta\beta(0\nu)$ is changed by the same factor of ten, the upper limit on neutrino mass would change from 1 to 3 eV.

Another recently proposed neutrinoless decay mode is the decay into two electrons and a massless neutral particle termed the majoron. This is predicted in several models in which lepton number is not conserved. Recently, a group from the University of South Carolina and Pacific

Northwest Laboratory (USC-PNL) reported⁹ the observation of double β -decay with majoron emission.

As the existence of the majoron would represent a new particle which could not be accommodated in the standard model, these reports attracted great attention. There are now three more experiments that have measured limits for neutrinoless β -decay with majoron emission. Two of these, one by a group from the University of California at Santa Barbara and Lawrence Berkeley Laboratory¹⁰, the other by a collaboration of the California Institute of Technology, the Swiss Nuclear Energy Institute and the University of Neuchâtel, use the same isotope, ^{76}Ge , used by USC-PNL. They set lower limits on the half life for this decay which rule out by a factor of 2 or more the rate originally indicated. More recently, Elliott *et al.*¹¹ published a limit on the same process in ^{82}Se which also excludes the earlier result.

Now that the two-neutrino double β -decay has been seen, experimentalists will

be trying to increase the sensitivity of their searches for neutrinoless double β -decay. Many theoretical attempts to go beyond the standard model of particle physics replace the massless Dirac neutrino with a Majorana neutrino of tiny mass. The only experimental constraint on such theories comes from the absence of neutrinoless double β -decay. □

1. Elliott, S.R., Hahn, A.A. & Moe, M.K. *Phys. Rev. Lett.* **59**, 2020–2023 (1987).
2. Furry, W.H. *Phys. Rev.* **56**, 1184–1186 (1939).
3. Caldwell, D.O. *et al.* *Phys. Rev. Lett.* **419**, 419–422 (1987).
4. Haxton, W.C. & Stephenson, G.J. Jr. *Prog. particle nucl. Phys.* **12**, 409–479 (1984).
5. Grotz, K. & Klapdor, H.V. *Phys. Lett.* **157B**, 242 (1985).
6. Doi, M., Kotani, T. & Takasugi, E. *Prog. theor. Phys. Suppl.* **83**, 1–172 (1985).
7. Kirsten, T. in *Nuclear Beta Decays and Neutrinos* (eds Kotani, T. *et al.* 81–92 (World Scientific, Singapore, 1986).
8. Manuel, O.K. in *Nuclear Beta Decays and Neutrinos* (eds Kotani, T. *et al.* 71–80 (World Scientific, Singapore, 1986).
9. Avignone, F.T. III *et al.* in *Proc. Salt Lake City Meeting* (eds DeTar, C. & Ball, J.) 359–369 (World Scientific, Singapore, 1987).
10. Fisher, P. *et al.* *Phys. Lett.* **192B**, 460–462 (1987).
11. Elliott, S.R., Hahn, A.A. & Moe, M.K. *Phys. Rev. Lett.* **59**, 1649–1651 (1987).

Michael S. Witherell is in the Department of Physics, University of California, Santa Barbara, California 93106, USA.

Regulation of transcription

Fraudulent promotion

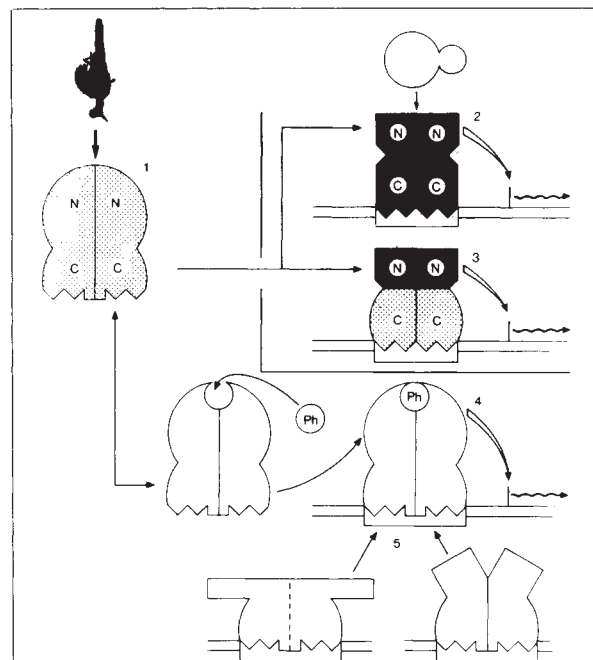
N.J. Short

IN AN earlier News and Views article¹, I discussed the possibility that the oncoprotein JUN represents a corrupted form of the transcription factor AP-1. The evidence for this, though quite compelling,

was entirely circumstantial, and more direct proof was needed to put the matter beyond doubt. Since then, new work² by Robert Tjian and colleagues has settled the issue unequivocally, while at the same

time affording a fascinating glimpse of just how complex some systems for transcriptional regulation may turn out to be.

As I related previously¹, the first hint that something unsuspected was afoot came from the detection of a similarity between the carboxy-terminal end of JUN and a yeast transcription factor, GCN4. This resemblance was not merely fortuitous; the carboxy-terminal end of JUN was soon found to be capable of replacing its counterpart in GCN4, implying that the two proteins must bind to very similar sequences. The sequence bound by GCN4 was known, however, and bears a remarkable resemblance to a motif found in mammalian promoters and enhancers. A protein which recognizes this motif, AP-1, had recently been purified and shown to mediate increased levels of transcription in



An evolving picture of the relationships between proteins activating transcription. The carboxy-terminal domain of JUN can activate transcription in yeast as shown in the figure from my earlier article (1–3), but it is now clear that the cellular homologue of JUN (4) is the most abundant of several proteins which recognize the AP-1 binding site (5).

response to treatment with phorbol ester. It was therefore natural to conjecture that the cellular homologue of JUN might be AP-1.

Bohmann *et al.* in their new work² show that this is indeed the case. Antibodies to JUN crossreact with AP-1. The cellular homologue of the *jun* gene, *c-jun*, has been isolated from a human genomic library and shown to encode six peptides which form part of AP-1. Finally, a complementary DNA version of *c-jun*, when expressed in bacteria, produces a protein with DNA-binding properties virtually indistinguishable from those of AP-1. The conclusion that *c-jun* really is the gene encoding AP-1 seems unavoidable. On the other hand, the differences between *jun* and *c-jun* are considerable, and suggest that JUN is not impersonating a simple AP-1 so much as an AP-1 that has bound phorbol esters. If so, this would confirm the earlier idea¹ that the cell hosting JUN might be permanently subject to some of the effects of these tumour promoters. A forthcoming report from Michael Karin's group³ (which also confirms much of the work of Bohmann *et al.*) supports this idea by showing that JUN activates transcription strongly in transfected cells without the need for phorbol esters.

Although this is significant in itself, some of the avenues of enquiry it suggests could be even more interesting. One of these is whether there may not be other proteins that bind to the AP-1 site. Bohmann *et al.*² mention that they have found at least one other gene with homology to *jun*, and that another group has found a third such gene in mice. Moreover, when AP-1 is purified on affinity columns bearing its binding site, the result is a preparation containing one major and three minor polypeptides. It therefore seems probable that JUN belongs to a multigene family, all the members of which have very similar DNA-binding domains and recognize the same sequence. The common DNA-binding domain plainly has a very long evolutionary history in this case, as it can still replace its yeast counterpart in GCN4. Despite this, the remaining parts of JUN and GCN4 bear no resemblance to each other, and there is no counterpart in JUN for the acidic amino acids that activate yeast transcription⁴. This evolutionary conservatism thus appears to be a case of Walter Gilbert's "exon shuffling"⁵, resembling the relationship between *cdc10* in fission yeast and *Notch* in fruitflies (which share a domain but not, so far as is known, any function⁶) rather than the relationship between the *cdc2s* in fission yeast and humans (which share both structure and function⁷). DNA-binding domains seem to have a particular predelection for being shuffled; this characteristic has been used, for instance, to isolate several genes from mice, all pos-

SN1987A pulsar slow to reveal itself

THE light curve of supernova 1987A has been falling now for several months, following an exponential law, with a half-life of 78 days. This is convincing evidence that the power behind the luminosity is the radioactive decay of ⁵⁶Co to ⁵⁶Fe; the ⁵⁶Co comes from the rapid decay of ⁵⁶Ni, about half a solar mass of which was synthesized in the original explosion. At the recent meeting of the American Astronomical Society in Austin, Texas (10–15 January 1988), direct evidence for the radioactive decay model was presented in the detection by G. Share *et al.* (reported in full on page 416 of this issue) of characteristic γ -ray lines of ⁵⁶Co. Moreover, the light curve is beginning to show signs of a slight turn-down below the exponential curve, as expected: when γ -rays from the decay start to leak through the thinning shell, less energy is available to be degraded into optical emission.

Nucleosynthesisists in Austin were evidently pleased that SN1987A is following their calculations with such admirable precision. But there are a few puzzles. The brightness in some parts of the ultraviolet is increasing, and IAU circulars 4530 and 4532 report increases in X-ray emission in the 6–28 keV energy range.

Advocates of the pure-radioactivity model attribute the ultraviolet emission to fluorescence of a shell of circumstellar material. This shell would be the result of mass loss by the supernova progenitor in an earlier red-giant phase, and radiation from the explosion will only just reach it to excite the fluorescence.

The same advocates generally ascribe the observed soft X-rays to experimental error; more particularly to contamination of the observations by the powerful source LMC X-1, not far in the sky from SN1987A. But another explanation is that the various increases in emission signal the emergence of a second power source — a pulsar.

Everyone agrees that a neutron star is lurking within the supernova remnant, and if it is spinning and has a magnetic field it

must transfer energy to the nebula. The mechanism and efficiency of this transfer are hard to guess at this early stage of development, but as radioactive decay declines and is replaced by a constant energy supply from the pulsar, the light curve should eventually level out.

There is no sign yet of this happening, and the light level is now comparable to the output of the Crab Nebula, which is entirely powered by its central pulsar. If the decline continues, pulsar theorists will begin to worry, although the birth of pulsars is a complex and poorly understood process, so that escape clauses can be found. It may be, for example, that pulsars are not born with a strong magnetic field but generate one by an internal dynamo mechanism, so that the pulsar will gradually turn on as the months pass.

Observers are also looking out for more specific signs of the presence of a pulsar. Good evidence, according to Alice Harding of NASA/Greenbelt, would be the detection of ultra-high-energy γ -rays, perhaps up to 10¹⁵eV, from the supernova remnant. The spinning magnetic field of the pulsar is able to accelerate charged particles, mostly electrons and protons, to these enormous energies; the protons collide with atoms in the supernova ejecta, creating, among other things, neutral pions which decay rapidly into γ -ray photons, some of which are able to escape the nebula and reach us. This is a lengthy chain of events, but it is widely accepted that only a pulsar is capable of generating such huge energies in individual particles, and their detection by atmospheric air showers or Cerenkov radiation on Earth would be proof for most astronomers that a pulsar exists.

Concern over the absence of a direct sighting of the pulsar has not yet turned to embarrassment, but the next few weeks and months, as energy from radioactivity falls off, will be crucial. A prolonged silence may cause even the most stalwart friends of pulsars to think about revising their ideas.

David Lindley

sessing the so-called 'zinc finger' DNA-binding domain and so presumably involved in transcriptional regulation⁸. Nevertheless, no proteins other than the AP-1s are yet known to share the GCN4-binding domain, and it remains to be seen whether this element can be modified so as to recognize different DNA sequences in different proteins.

Whether or not it can, the idea that there is a family of AP-1s makes the transcriptional response to phorbol esters and its relationship to carcinogenesis seem even more complicated than it did already. The effect of phorbol esters on the other AP-1s is not yet known, but at least four further transcription factors have been described whose effectiveness

is increased by these compounds⁹. Are any of these in fact the products of multigene families? How do the factors in this vast array interact with one another? And which, if any, of them are potential oncogenes like *c-jun*? There are plainly many interesting answers to look forward to. □

1. Short, N. J. *Nature* **330**, 209–210 (1987).
2. Bohmann, D. *et al.* *Science* **238**, 1386–1392 (1987).
3. Angel, P. *et al.* *Nature* (in the press).
4. Ma, J. & Ptashne, M. *Cell* **51**, 113–119 (1987).
5. Gilbert, W. *Science* **228**, 823–824 (1985).
6. Breeden, L. & Nasmyth, K. *Nature* **329**, 651–654 (1987).
7. Lee, M. G. & Nurse, P. *Nature* **327**, 31–35 (1987).
8. Chowdhury, K. *et al.* *Cell* **48**, 771–778 (1987).
9. Chiu, R. *et al.* *Nature* **329**, 648–651 (1987).

N. J. Short is in the Department of Biophysics, Cell and Molecular Biology, King's College London (KQC), 26–29 Drury Lane, London WC2B 5RL, UK.

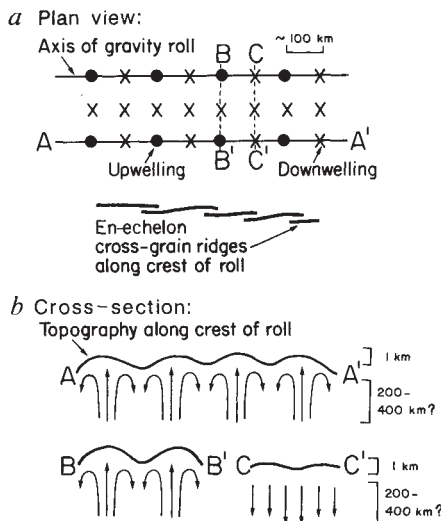
Cracks in the Pacific plate and mantle convection

SIR—Winterer and Sandwell¹ interpret 'cross-grain' ridges on the Pacific plate which are oblique to both spreading-centre parallel topography and fracture zones as large-scale tensional cracks up to 200 km long and spaced ~35 km apart. The cracks occur along topographic highs which exhibit gravity signals with wavelengths of ~200 km and a linear continuity of up to thousands of kilometres. These gravity 'rolls' may result from shallow-level convection in the upper mantle which has been sheared into rolls by the motion of the overlying Pacific plate².

Cross-grain ridges share two characteristics with overlapping spreading centres (OSCs) on the East Pacific Rise (EPR)^{3,4}, which reinforce Winterer and Sandwell's tensional crack interpretation and suggest an analogy between the two structures which further elucidates the possible connection between cross-grain ridges and mantle convection. First, both structures consist of ridges which overlap in an en-echelon fashion by a distance approximately three times their offset^{1,4}. This is true of many of the OSCs which have been carefully mapped, and is characteristic of tensional cracks ranging over ten orders of magnitude in size^{4,5}. Second, in both OSCs⁴ and cross-grain ridges¹ the summits of the overlapping en-echelon ridges plunge over long distances towards the region of overlap.

The first similarity supports Winterer and Sandwell's hypothesis that the cross-grain structures are large tensional cracks, explains why the overlap distance is three times the offset, and suggests an explanation for their en-echelon offset. Using the displacement discontinuity method, we have found^{6,7} that when a tensional stress is applied perpendicular to two parallel cracks, they propagate towards each other until the ratio of overlap distance to offset distance attains a value of ~3. For overlap to offset ratios greater than ~3, the crack propagation force decreases to a very small value⁶, and propagation ceases. Our calculations also show that unless cracks are perfectly colinear (which is not possible in nature), en-echelon offsets will form, because the first interaction between the approaching crack tips is a deflection away from each other (see ref. 5, Fig. 8). A uniform 3:1 ratio of crack overlap to offset thus implies that the two cracks have been subjected to deviatoric tension oriented perpendicular to their length, that they have been active at the same time for at least part of their respective histories, that the cracks have propagated towards each other, and that propagation ceased when the ratio of overlap to offset reached ~3.

Now consider the similarity in axial depth profiles between the EPR and the



a, Plan view showing possible relationship between mantle convection, gravity rolls and cross-grain ridges. b, Cross-sectional views. If the analogy between overlapping spreading centres on mid-ocean ridges and cross-grain ridges is valid, then shallow places along the gravity rolls may overlie regions of mantle upwelling, whereas deep areas and the locations where en-echelon cross-grain ridges overlap may overlie regions of downwelling.

cross-grain ridges; in both cases, the depths steadily increase near the tips of the en-echelon ridges (compare Fig. 4 in ref. 1 with Fig. 3 in ref. 4). For the EPR, several lines of evidence suggest that this steady increase in depth near OSCs is due to shallow convective upwelling in the upper mantle beneath high portions of the ridge. Upwelling results in decompression melting and segregation of partial melt as it migrates upward, and to flow of partial melt downhill along the strike of the ridge⁴.

If the analogy between cross-grain ridges and OSCs holds, then the following hypothesis can be considered. Shallow portions of cross-grain ridges correspond to sites of short-wavelength upwelling of asthenosphere in a three-dimensional framework of mantle convection (see figure). Upwelling leads to doming and cracking of the overlying lithosphere. Partial melt segregates and rises, then migrates perpendicular to the axis of deviatoric tension (parallel to the axes of the cross-grain ridges), creating a graded and plunging profile away from the site of upwelling. Conversely, deep portions of cross-grain ridges, where en-echelon overlap occurs, correspond to sites of downwelling of asthenosphere and are relatively starved of partial melt (see figure). Bathymetric swath mapping coupled with seismic reflection and refraction measurements have supported a similar hypothesis for the EPR and similar experiments could support or refute the hypothesis offered here.

The analogy between en-echelon over-

lapping of cross-grain ridges and OSCs is imperfect at best because cross-grain ridges are not spreading and the sizes of the structures are quite different. But the characteristic geometry of OSCs and cross-grain ridges should not depend on spreading or on size; the relationships hold for cracks in glass on a microscopic scale and for dykes in areas that are not spreading⁴.

KEN C. MACDONALD

Department of Geological Sciences,
University of California,
Santa Barbara,
California 93106, USA

1. Winterer, E.L. & Sandwell, D.T. *Nature* **329**, 534-537 (1987).
2. Haxby, W.F. & Weissel, J.K. *J. geophys. Res.* **91**, 3507-3520 (1986).
3. Macdonald, K.C. & Fox, P.J. *Nature* **302**, 55-58 (1983).
4. Macdonald, K.C., Sempere, J.-C. & Fox, P.J. *J. geophys. Res.* **89**, 6049-6069; 6301-6305 (1984).
5. Sempere, J.-C. & Macdonald, K.C. *Tectonics* **5**, 151-163 (1986).
6. Macdonald, K.C., Sempere, J.-C., Fox, P.J. & Tyce, R. *Geology* **15**, 991-1086 (1987).
7. Pollard, D.D. & Aydin, A. *J. geophys. Res.* **89**, 10017-10029 (1984).

Periodic extinctions within the Cenozoic

SIR—I was happy to read the letter "Is the periodicity of extinctions a taxonomic artefact?" by Patterson and Smith¹ together with the reply by Sepkoski². Having no expertise with either the echinoderms or fishes discussed by Patterson and Smith, nor with the statistical techniques that gave rise to the matter, I am nevertheless concerned that there is much more to the question. Namely, if the alleged 26 Myr extinction periodicity is to be taken seriously one should find good evidence for it within the Cenozoic. The Cenozoic is far better known biostratigraphically and palaeontologically than any other comparable interval since the opening of the Cambrian. Sepkoski and Raup³ (Table 3) allege a major, 'significant' extinction event at the Eocene-Oligocene boundary, although biostratigraphers, to whom Sepkoski paid lip-service in his reply, have never recognized this as a major boundary in the same terms as the Cretaceous-Tertiary, Permian-Triassic, Triassic-Jurassic, or any of the other chief extinctions well known to biostratigraphers since the middle of the last century. Even if one accepts the Eocene-Oligocene boundary as a major, but poorly defined to date, extinction event one is still left with the problem of trying to recognize a major, 'significant' extinction event within the Miocene. No biostratigrapher of the past century and more has ever suggested the existence of a major boundary, one that could be interpreted as a global extinction event, within the Miocene. On these grounds alone the alleged 26 Myr periodicity fails. Additionally, in my own experience within the Silurian-

Devonian interval, an interval approximating 100 Myr, there is good evidence for a major extinction at the end of the Ordovician, and for another within the Upper Devonian, but no major extinction of any sort within the remainder of this very lengthy interval which exceeds 26 Myr by at least three times.

If the alleged 26-Myr periodicity is defined so as to include second, third, or even fourth order extinction events, defined in one way or another, then we are faced with the problem posed by data such as that summarized by House⁴, who showed for the Devonian that there are a very large number of minor extinction events occurring at intervals far below the 26 Myr level. This being the case, if one wishes to opt for shorter time intervals, one could favour just about anything between 5 and 26 Myr in view of the varied physical and biological uncertainties attending such work, with no great difficulty at all.

Smith and Patterson, and Sepkoski both allude to generic compilations prepared by Sepkoski⁵ as evidence in the extinction question. Through the courtesy of Sepkoski I was provided with a copy of his printout of the generic data for the brachiopods — a significant part of the Palaeozoic database. I am personally familiar with the Silurian–Devonian genera, and have contributed a reasonable fraction of the data in one way or another. Inspection of Sepkoski's data indicates that he does not understand the generic concept as used by most invertebrate palaeontologists. He has taken the known, upper time-range limit of each brachiopod genus as an extinction. Anyone familiar firsthand with the making, describing and revising of genera is certainly well aware of the gradational, transitional nature of many genera. This is particularly true for the larger genera, those which happen also to be commonly more abundant as individual specimens, and to be more cosmopolitan and eurytopic in the sense of occurring in more major environments and communities. After considering Sepkoski's brachiopod data for the Silurian–Devonian, a substantial fraction of his total brachiopod data, I estimate that at least 75 per cent of these genera could be considered by many taxonomists to have gradational relations with either ancestral or descendant genera, or with both. That is, they are paraphyletic in Patterson and Smith's terms. In other words, the bulk of his genus-level extinctions are mere name changing, book-keeping extinctions, not true extinctions.

In view of this situation I conclude that his generic compilations are of such limited value as to be of no use for recognizing any but the most obvious, nineteenth century type, extinction events, and to be of no value for recognizing

minor extinctions. In fact, Sepkoski used such generic data to recognize an alleged extinction event at the Pennsylvanian–Permian boundary. Experienced biostratigraphers and palaeontologists have long known that there is no positive evidence for such an extinction event.

It seems to me that the recognition, and magnitude estimation, of extinction events large and small could be done far better initially by paying more careful attention to the nature of the breaks within the well known, carefully tested biostratigraphic framework, than by using ever more sophisticated statistical devices⁶ with data of questionable quality.

ARTHUR J. BOUCOT

Department of Zoology,
Oregon State University,
Corvallis,
Oregon 97331, USA

1. Patterson, C. & Smith A.B. *Nature* **330**, 248–251 (1987).
2. Sepkoski, J.J. *Nature* **330**, 251–252 (1987).
3. Sepkoski, J.J. & Raup, D.M. in *Dynamics of Extinction* (ed. Elliot, D.K.) 3–36 (Wiley, New York, 1986).
4. House, M.R. *Nature* **313**, 17–22 (1985).
5. Sepkoski, J.J. in *Global Bio-Events* (ed. Walliser, O.H.) 47–61 (Springer, Berlin, Heidelberg, 1986).
6. Stigler, S.M. & Wagner, M.J. *Science* **238**, 940–945 (1987).

Founder effects and endangered species

SIR—Diamond and Rotter point out¹ that epidemiological studies of human genetic disorders demonstrate the significance of the founder effect. The examples they cite demonstrate that, for human populations begun from relatively few original founders and genetically isolated for many generations, unique or rare genetic disorders are not an uncommon consequence. Genetic disease and genetic load are the results of mutational events for which there is a common basis in all species. This is the theoretical underpinning for the search for animal models of human disease, for example.

Captive populations are now being established for many endangered species, and for many of these species founder numbers are very small. Often these populations are considerably smaller than for human populations. Additionally, in zoo management, unlike human populations, inbreeding between close relatives is not uncommon.

Several species now held in captivity have captive pedigrees that are complete from the founders to their descendants over many generations. Przewalski's horse, for example, is thought to be extinct in the wild. The entire world population of 660 animals, bred for 12 generations, traces its ancestry back to 13 individuals, and nearly one-quarter of all genes in these horses could be identical by descent from individual founders. The golden lion tamarin has been bred for six generations in captivity and the total world population of about 500 animals is

derived — with unequal representation — from 51 founders². The red ruffed lemur has been bred in captivity for seven generations. To date, the 125 living animals trace their ancestry to seven founders, three of whom contribute more than 70 per cent of the genes in the current population (O.A.R. and D. Brockman, in preparation).

Genetic diseases may already be present in all these species. In each one there are congenital anomalies or other defects that seem to be familial. The high degree of consanguinity in these pedigrees, as well as the difficulty in excluding environmental influences, makes straightforward diagnosis of genetic origin difficult.

In Przewalski's horse, there may be a familial vitamin E malabsorption syndrome³. There is also a coat-colour anomaly whereby affected individuals lack black pigmentation on their mane, tail and the extremities of their limbs. There is concern about an hereditary diaphragmatic hernia in the golden lion tamarin². In the red ruffed lemur, hairless offspring have been produced from matings inbred through a single founder individual⁴. Here, the mode of inheritance seems to be autosomal recessive. Also in red ruffed lemurs, funnel chest (*pectus excavatum*) has been described⁵.

These three species are remarkable only in that there are excellent pedigree data for many generations. It is possible that, as the number of generations in captivity increases, species whose captive populations are based on small numbers of founders will manifest genetically based deleterious conditions. Przewalski's horse, California condors and black-footed ferrets are now found only in captivity, but reproduction has only just begun in the ferrets and has not yet started for the condor.

When genetic diseases occur in captive populations of endangered species, management would benefit from carrier tests and prenatal diagnosis, as are used for some human genetic diseases caused by single gene defects⁶. Construction of such tests in endangered species requires approximately the same level of understanding of the species' genetic map as is necessary in humans, a goal far from realization.

OLIVER A. RYDER

Center for Reproduction of Endangered Species,
Zoological Society of San Diego,
PO Box 551, San Diego,
California 92112,
USA

1. Diamond, J.M. & Rotter, J.I. *Nature* **329**, 105–106 (1987).
2. Ballou, J.D. *International Studbook: Golden Lion Tamarin* (National Zoological Park, Washington DC, 1986).
3. Liu, S.-K., Dolensek, E.P., Adams, C.P. & Tappe, J.P. *J. Am. vet. Med. Ass.* **183**, 11–14 (1983).
4. Goodwin, L.G. *J. Zool.* **190**, 555 (1980).
5. Benirschke, K. *Int. Symp. Erkrank. Zootiere* **22**, 169–172 (1980).
6. Cooper, D.N. & Schmidtke, J. *Lancet*, 273 (1987).

St George and the dragons

Walter Gratzer

Free Radical: Albert Szent-Györgyi and the Battle Over Vitamin C. By Ralph W. Moss. Paragon House, New York:1987. Pp.316. \$22.95.

ALEXANDER Korda is said to have kept on his desk a card bearing the inscription "It is not sufficient just to be Hungarian". Nevertheless the Hungarians — or as some thought the Martians, who had established themselves with their unaccountable language and culture in a remote eastern European fastness, thence to colonize the Western world — took possession not only of the film industry but of a great deal of American science as well.

Perhaps the most flamboyant of these invaders was Albert Szent-Györgyi, who arrived in 1947 with his group of myrmidons, a Nobel prize and a mixed reputation. Szent-Györgyi had an abundance of vitality and charisma; in the previous 50 years he had experienced adventure and adversity, loyalty and betrayal, success and calamity. His nature was full of contradictions, of overwhelming charm and generosity, intermingled with egotism and vindictiveness. He was a brilliant and outrageous ham, as no one who heard him lecture is likely to forget. He was four times married, twice to women 50 years his junior, he bore himself heroically in two world wars, could, had he wished, have been a Head of State, and he might have had a share of three Nobel prizes, but the one he did get left a legacy of rancour and bitterness. He was loved, adulated and by some detested. He is in short a rewarding subject for a biographer and Ralph Moss, who came to know his man as a friend, has made the most of his opportunity. He has written an absorbing book, clear-eyed, scholarly and scrupulous (and far better than the tiresome title might suggest).

Albert Szent-Györgyi was born into the minor aristocracy and had a distinguished intellectual lineage on his mother's side. An uncle, who was a professor in the Budapest medical faculty, drove him into the arcane discipline of proctology, possibly as a mark of his low esteem, but also that Albert might minister to the piles to which he was a martyr. This unpropitious career was interrupted by the First World War, in which the luck that was to be his constant companion for the next 30 years first showed itself. Szent-Györgyi was decorated and lived through the annihilation of the Hungarian army in the Ukraine and the ravages of cholera,

smallpox and typhus that wiped out most of the survivors. Tiring of the futile carnage, he shot himself through the arm, was invalided home and, having recovered, narrowly escaped a posting to the north of Italy, where an even more desperate campaign was in progress.

After the war there followed the *Wanderjahre*, spent in penury with his wife and young daughter in a succession of European countries. Szent-Györgyi had the good fortune to rub up against some of



Vital observation — Szent-Györgyi at work at the Woods Hole Institute with his cousin Andrew.

the great founders of the new science of biochemistry, notably Michaelis in Berlin and Gowland Hopkins in Cambridge. This period saw the beginning of his lifelong interest in biological oxidation, then a subject of furious debate between the two titans of German biochemistry, Wieland and Warburg. Szent-Györgyi's observation, made in total isolation in a physiology department in Gröningen, that carboxylic acids were implicated in electron transfer, prefigured the discovery of the citric acid cycle. Starting from a totally false premise, he next demonstrated the presence of a strong reducing agent — which he surmised for no obvious reason to be identical with the principle that had been shown to exist in citrus fruits — in adrenal glands. This substance, which he wanted first to call ignose and then godnose, in the mistaken belief that it was a sugar, was termed hexuronic acid at the insistence of the editor of the *Biochemical Journal*. It turned out of course to be ascorbic acid.

Back in Hungary, with his own research group at last, thanks to the good offices of Horthy's enlightened Minister of Education, who became his first great patron, he had his next big stroke of luck. There arrived unbidden in his laboratory in Szeged a young Hungarian American, Joseph Svirbely, who had worked in Pittsburgh in one of the several centres then competing to isolate vitamin C. Svirbely knew how to assay for antiscorbutic activity and at once established that hexuronic acid was indeed vitamin C. A letter was sent to Svirbely's former professor, C.G. King, who promptly and shamelessly jumped the gun and got a paper off to *Science*. Szent-Györgyi next found, quite fortuitously, that paprika, which happened to be the principal crop of the Szeged district, was an exceptionally abundant source of vitamin C, and many grams of crystallized material were soon isolated and sent to Haworth in Birmingham for structure determination.

From the fog of controversy that quickly gathered around the discovery of the vitamin, Szent-Györgyi emerged with the Nobel prize and King — so far at least as the United States was concerned — with the credit. The American press accused Szent-Györgyi of defrauding their hero and traducing his character. Szent-Györgyi could never forget the affair — the battle of Moss's title — and apparently acceded with eagerness to Moss's request to write his biography, mainly so that the record could finally be set straight. Svirbely, caught in the middle, gave up science and when, as a tremulous septuagenarian, he was interviewed by Moss, was still too cowed to want to talk about it.

In 1939 Szent-Györgyi turned his attention to muscle, and he and his group discovered myosin and actin and the relaxation of the rigor state by ATP, thereby initiating an entirely new field of research.

Szent-Györgyi became Rector of the preposterously renamed Royal Nicholas Horthy University in Szeged at a time of encroaching racism and intolerance; he bravely resisted the fascists, who had captured control of the universities, and created an oasis of liberal enlightenment. With the engulfment of Hungary by the Germans, Szent-Györgyi went underground and began a dangerous game of hide-and-seek with the Gestapo, a beard his only disguise. His luck held, and when the German war effort began to crumble he undertook clandestine negotiations with the Allies on behalf of the chastened Horthy regime.

In 1945 he surfaced as a national hero;

turning down the presidency of the new Hungarian Republic, he applied himself instead, with the help of the Russians, over whom he had acquired considerable influence, to establishing a laboratory in Budapest. Political quarrels and disillusion followed, his closest associates were denounced and dismissed, and Szent-Györgyi, suddenly a target of political obloquy, left for America.

The indifference of his reception there came as a shock; his hyperbolic style, and perhaps the memory of the vitamin C episode, exposed him to the malign machinations of power-mongers, such as



Woods Hole, 1952 — Marta Szent-Györgyi makes the tea, while Albert (left, seated) entertains his colleagues.

Detlev Bronk, and he also came under the baleful scrutiny of J. Edgar Hoover. Szent-Györgyi soon discovered that in the American academic community esteem is a highly perishable commodity (this enduring truth was memorably encapsulated in a *New Yorker* cartoon: "So he discovered fire and the wheel" says one caveman to another, gazing at a third, "but what has he done since?"). He reacted by becoming increasingly arrogant and dogmatic in his utterances. Yet the charm still worked; rejected by the funding agencies, he managed to persuade the meat barons of the Armour Company that their trade was muscle and that investment in his research would yield commercial treasure. The Institute of Muscle Research, which he set up at Woods Hole, did not in fact continue long with muscle research (my recollection is that, by his own account, he gave up working on muscle when it became known that myosin and actin were in separate filaments). Instead he returned to certain early and inchoate ideas about vitamin C and biological oxidation, and became persuaded that in them lay the key to cancer.

At this time Szent-Györgyi also became preoccupied with a vision of a new 'sub-molecular biology', and of the importance

in biology of excited-state phenomena, such as charge transfer complexes. The requirement that was laid on him to write grant applications he considered as an affront not only to his professional standing but to logic as well, for if one knew and could describe what one was going to discover, what would be the point in discovering it? The American cancer establishment was not impressed and Szent-Györgyi reciprocated with a fine show of contempt for their ways; cancer, he once said, was coming close to keeping more people alive than it killed. Nevertheless, such was the allure of his personality that until well into his ninetieth decade he did not want for patrons or scientific followers, whose talents he often perverted.

Alas, the edifice proved to have been built of straw; Szent-Györgyi's pronouncements about the imminence of an explanation and cure for cancer became increasingly extravagant and irresponsible.

The advances in molecular biology, that were opening up cancer research, passed him by and in the end it became clear that nothing remained but an old man's unreasoning obsession. The foundation, set up to support his cancer research by a philanthropic lawyer, wholly in his thrall, collapsed in squalor and recrimination; influenced by his young wife, Szent-Györgyi became alienated from his family and friends. He had become a relic, a *monstre sacré*, lonely embittered and unfulfilled. He died at 93, still pursuing his star, for, as he had said, he knew no other way to live. His was the true tragedy of old age, which Oscar Wilde put into the mouth of Lord Henry Wotton in *The Picture of Dorian Gray* — not that one is old, but that one is young. □

Walter Gratzer is in the Medical Research Council Cell Biophysics Unit, King's College London (KQC), 26-29 Drury Lane, London WC2B 5RL, UK.

In the balance

Paul Langacker

Physics of Massive Neutrinos. By Felix Boehm and Petr Vogel. Cambridge University Press: 1987. Pp.211. £25, \$34.50.

ONE of the most intriguing questions in particle physics is whether the neutrino has mass. Experimentally, any neutrino mass must be tiny, and the highly successful standard electroweak model predicts massless neutrinos. However, it is clear that the standard model is not the ultimate theory of nature. Most theories of the new physics that must exist below 10^{-16} cm predict a non-zero mass at some level, and often that it is of the peculiar Majorana character, which means that there is no fundamental distinction between neutrinos and antineutrinos.

Neutrino mass is also of vital importance for astrophysicists. The Universe is presumably filled with enormous numbers of neutrinos left over from the first second of the Big Bang. If these have even a small mass, in the $10 \text{ eV } c^{-2}$ range, they would dominate the energy density of the Universe and account for the missing (dark) matter. Furthermore, coherent interactions of very light ($\leq 10^{-2} \text{ eV } c^{-2}$) neutrinos in the Sun could account for the missing solar neutrinos in the famous ^{37}Cl experiment.

The subject is of bewildering complexity, however. Experiments and the theoretical ideas needed to interpret them overlap particle, nuclear and astrophysics, and the field is confused by conflicting experiments and calculations. Boehm and Vogel's little book goes a long way towards clarifying these difficult issues. It provides a readable and concise, but

still fairly thorough, guide to most aspects of the field.

The formalism of Majorana and Dirac neutrinos, theoretical models and neutrino interactions are briefly described. There is an excellent treatment of the various (conflicting) tritium beta-decay experiments, an overview of the varied experiments searching for small mixings with possible heavy neutrinos, and thorough accounts of all aspects of double beta-decay and neutrino oscillations (two subjects on which the authors have worked extensively). There are also short discussions of cosmological implications and the impact of Supernova 1987A.

The book provides a balanced treatment of conflicting experiments. It is generally accurate and complete, though there are occasional lapses (such as wrong attribution of the Gelmini-Roncadelli Majoron model and omission of any mention of the suggestive CERN PS-191 oscillation results). Perhaps its main weakness is in the treatment of the solar neutrino problem. A discussion of the several parameter ranges relevant to matter-induced oscillations and how they could be distinguished from alternative models would have been useful.

Overall, this is an excellent introduction and guide to an important and complex field. □

Paul Langacker, currently at DESY, Notkestrasse 85, D-2000 Hamburg 52, FRG, is a Professor of Physics at the University of Pennsylvania.

New in reference

- *Information Resources in Toxicology*, 2nd edn, by P. Wexler. (Elsevier, Dfl. 195, \$85.)
- *Dictionary of Genetics and Cell Biology* by N. Maclean. (Macmillan, hbk £27.50, pbk £10.95.) In the US, New York University Press.
- *Dictionary of Microbiology and Molecular Biology*, 2nd edn, by P. Singleton and D. Sainsbury. (Wiley, £69, \$150.)

Completely logical thinking

Martin Davis

Reflections on Kurt Gödel. By Hao Wang. MIT Press: 1987. Pp. 336. \$25, £22.50.

THIRTY-FIVE years ago, newly arrived in Princeton, New Jersey, I was driving towards the Institute for Advanced Study where I was to be a temporary member, when I found my way blocked by an odd pair walking slowly and obliviously in front of my vehicle. The taller man was quite unkempt while the other was immaculately dressed in a business suit and carried a briefcase. As I cautiously passed them, I could see that it was Einstein and Gödel. "Einstein and his lawyer", my wife quipped.

It was not only in their dress that these good friends differed. In 1953, Einstein declared "You know, Gödel has really gone completely crazy. . . . He voted for Eisenhower". Einstein, a socialist, could not imagine voting for the Republican candidate.

Further, Einstein's special theory of relativity begins by challenging the *a priori* character of space and time. Gödel began his career as a student in Vienna in the 1920s, where the dominant philosophical tendency was that of the positivistic Wiener Kreis, and his outlook was very different. To him the abstract entities of mathematics — numbers, sets and functions — possessed an objective existence. Propositions about these entities were either true or false on the basis of their meaning, independent of the possibility of proving them in the formal logical systems of Bertrand Russell, Thoralf Skolem and David Hilbert. To members of the Kreis, 'truth' was to be 'explained' as meaning provability in one of these systems.

Gödel's viewpoint was certainly justified by the results: he was able quickly and decisively to solve a number of key problems about symbolic logic set by the great Hilbert. In his doctoral dissertation, he proved the completeness of the standard axiomatization of predicate logic, and by 1930 he had shown the incompleteness of the axioms of arithmetic, exhibiting a *true* sentence of arithmetic that could not be proved from these axioms.

Gödel's work was truly the fruit of his philosophical outlook. The methods needed to obtain his completeness theorem had been well known to Skolem and to Jacques Herbrand. It was only their philosophy that prevented them from drawing the conclusion. The methods Gödel introduced to prove the incompleteness of arithmetic were entirely novel, but not really difficult. However, for researchers in mathematical logic and the foundations

of mathematics, Gödel's work utterly transformed the intellectual landscape. In addition, it set the stage for the work of Alan Turing from which derives the concept of the all-purpose computer.

Gödel's incompleteness proof made use of the diagonal method introduced in the nineteenth century by Georg Cantor to show that, in a suitable sense, every collection has 'more' subsets than elements. Gödel used the diagonal method in a constructive manner: given an axiomatization of arithmetic, Gödel's proof could be used to produce explicitly a true sentence not provable from that axiomatization. Cantor had conjectured his 'generalized continuum hypothesis' (GCH): every collection of subsets of a given infinite set has the same number of elements as either the given set or as the set of *all* of its subsets. GCH was Gödel's next major project. Extending Russell's predicative hierarchy through all of Cantor's transfinite ordinal numbers, Gödel arrived at what he called the constructible sets. He was able to show that the constructible sets satisfy all the usual axioms of set theory, and that GCH, relativized to the constructible sets, is true. It followed that GCH could not be *disproved* from these axioms.

Gödel's work on GCH was completed by 1938. He was a member of the Institute for Advanced Study from 1940 until his death early in 1978. Except for brief excursions (including one into general relativity), Gödel's work during this 40-year period, which mostly remains unpublished, seems to have been mainly in philosophy.

Hao Wang, himself a distinguished logician, has performed a remarkable service in producing this book. He spent a great deal of time with Gödel, seeking to understand his highly unusual outlook, and here provides not only a wealth of biographical information about Gödel, but also Gödel's own explanation of how he came to his revolutionary discoveries. In addition, Wang initiates the reader into some of the more astonishing aspects of Gödel's belief system and gives his own views as well. Gödel produced a new proof that God exists and believed that a truly 'scientific' philosophy and a 'rational' religion could and should be developed. He thought that it would then be possible to settle decisively such questions as whether the human mind could have evolved. Gödel even thought it possible that Leibniz's project of a universal language of science in which all questions reduce to computations could be accomplished in a few years.

Hao Wang has provided us with an invaluable source book on one of the most remarkable thinkers of our century. □

Martin Davis is a Professor of Mathematics and Computer Science at the Courant Institute of New York University, 251 Mercer Street, New York, New York 10012, USA.

Seeking a solution

Paul Doty

Making Weapons, Talking Peace: A Physicist's Odyssey from Hiroshima to Geneva. By Herbert F. York. Basic Books: 1987. Pp. 346. \$22.95.

THE small group of scientists and engineers who brought the world into the Nuclear Age have nearly all departed. Only those who were very young at the creation or were slightly older and blessed with long life are with us. Thus there are only a few testaments to be filed before the record of the witnesses will be complete. Herbert York's autobiography is one of these last documents.

It is an engaging work: unpretentious and conversational in tone, with his daily routines and reflections claiming equal time with the inside description of the momentous parade of events which he participated in or observed close at hand. Although York came to the Nuclear Age a bit late, he is distinguished by having stayed with it throughout his life and by having played a great variety of roles.

York came from a railroading family in Rochester, New York, and attended the University of Rochester, receiving a Master's degree in May 1943 at the age of 21. He put aside further graduate work to join the Rad Lab at Berkeley and to be inspired by E. O. Lawrence. His first assignment was to a small group intent on extracting uranium-235 by the electromagnetic isotope-separation method. Although this prototype work went well, the mammoth production version being built in Tennessee did not. Within eight months York was transferred there, and through Herculean labour, with others directed by Lawrence, revised the production units so that by mid-1945 the required amount of U-235 was ready to fuel Little Boy, the bomb that destroyed Hiroshima.

Initially, York was elated at having had a part in ending the war and perhaps in making war obsolete. Subsequently his doubts grew, but in the end he came to believe that the deterrent effect of the bomb and the changes in historical behaviour that fear of the bomb has induced have saved us from world war so far and "maybe, just maybe [this situation will persist] until the time when we can achieve a real and permanent solution".

Making Weapons, Talking Peace records his personal role in buying time for this slow evolution to occur — by becoming the first Director of the Livermore Laboratory (at the age of 30); by leading the early US military effort in space; by wise involvement in most of the major technical defence decisions of the 1950s, and to a lesser extent in those of the 1960s; by encouraging early arms control efforts and



Herbert York — involvement in the future.

reaching out to meet the Soviets in this regard; by helping to build the University of California at San Diego; by returning to Washington in the Carter period, first to advise on defence matters and then as chief negotiator on the Comprehensive Test Ban; by entering again into public discussion of the nuclear future and founding his Institute on Global Conflict and Cooperation at San Diego, and rethinking how to influence the balance between maintaining the peace through the threat of mutual annihilation (which cannot be a permanent solution) and building a more secure alternative international relationship.

But the impact of the book is in its vignettes. For example, York brings to life the way the US approach to missiles developed in the mid-1950s through a number of small Defense Department Committees orchestrated by von Neumann. By the time von Neumann died (1955), the main elements of the US strategic missile programmes had been designed and development of them had begun. York is particularly persuasive in showing how these plans came to fruition despite the contrary public view due to the riveting success of Sputnik. The transfer of the 'outside expert committee' approach to broader issues, initiated by Eisenhower, faltered at first. York is unsparing in criticism of his own role in the Gaither Panel's (1957) alarmist assessment and in praising Eisenhower's good judgement in not implementing it. But with the appointment of James Killian to establish a new President's Science Advisory Committee (of which York was a member), the situation was redressed and planning for an innovative US space programme got under way in earnest.

In the summer of 1960 York suffered a serious heart attack. Not yet 40, he retreated to the Chancellorship at San Diego where he subsequently became Graduate Dean and then Institute

Director. During most of the next 20 years he maintained an active role in a number of Washington assignments. His recognition of the potential role of arms control grew. He backed the mobile Midgetman, with its much lowered yield over the MX, throughout this period. And he was instrumental in delaying the development of anti-satellite systems, ending up as the Defense Department representative on negotiating a ban on them; the negotiations fell victim to the Soviet invasion of Afghanistan, but may have contributed to the very slow pace of development of such systems on both sides.

Early in 1979, before the anti-satellite negotiations collapsed, York was appointed chief US negotiator at the Comprehensive Test Ban talks. The timing was equally bad, but with perseverance much was accomplished during the year that remained before the election of Reagan and termination of the negotiations. York's detailed discussion of this period is among the most instructive parts of the book.

In the closing chapter York picks up his earlier involvement in the public consideration of arms control, recounting his

meetings with the Pope and the Catholic bishops and giving his current reflections on the subject. He dwells effectively on his period of pessimism during the Vietnam War (in 1970 he wrote in *Life* magazine that "Unless we nerve ourselves to make the attempt [to roll back the arms race] and make it soon, we are quite simply doomed"). But by the end of this long account of his personal odyssey he has become more optimistic again:

... notions that we must 'do something' radical about either the 'Soviet Threat' or the 'Nuclear Arms Race' very soon or be doomed have been with us for more than forty years. So far they have always proved to be wrong, and I expect they will remain so for the foreseeable future ... If we are wise enough, we will use the [time which a balance of power is buying us] to find a way out of the grand nuclear dilemma.

This autobiography sets the standards of involvement that the post-nuclear generations will need if a way out is to be found. □

Paul Doty is Mallinckrodt Professor of Biochemistry at Harvard University, and Director Emeritus of the Center for Science and International Affairs, John F. Kennedy School of Government, Cambridge, Massachusetts 02138, USA.

Manual work

J. de Belleruche

Neurochemistry: A Practical Approach. Edited by A.J. Turner and H.S. Bachelard. IRL Press: 1987. Pp.227. Hbk £26, \$47; pbk £17, \$32.

AN INVALUABLE book in the early days of neurochemistry was *Practical Neurochemistry*, by McIlwain and Rodnight, but since then there has been a reluctance on the part of researchers to write critically and instructively about methodological developments in the subject. This volume goes a long way towards filling the gap.

Nine approaches widely used in studies of the nervous system are discussed and assessed, and specific practical protocols are presented. The topics covered are subcellular fractionation, cell culture, immunocytochemistry, bioluminescence, receptors, protein phosphorylation, phosphoinositide turnover, isolation of cell nuclei and expression of receptor mRNA in *Xenopus* oocytes.

The general format set by the editors is excellent — the procedures are laid out in clear steps and the information included is comprehensive in scope, with valuable notes about 'best buys'. This format still allows a measure of variety. Some chapters are arranged strictly as practical protocols (for example that on subcellular fractionation), while others (such as the chapter on neuroreceptors) are in a highly readable form and explain the pitfalls that

have been encountered in the development of particular techniques.

Other contributions, such as those on bioluminescence and protein phosphorylation, provide the meticulous detail needed to translate the information into practice and thereby save many hours of trial and error in finding the optimal conditions or best source of materials. Another attractive feature of these chapters is the illustration of the methods they describe with typical experimental results. The accounts of cell culture and immunocytochemistry also provide a comprehensive background for beginners embarking on work using these approaches.

Techniques such as *Xenopus* oocyte translation of exogenous mRNA have made a fundamental contribution to our understanding of the nature of oligomeric receptor proteins, notably the nicotinic receptor. The account of the properties of *Xenopus* that make this possible, along with further applications of the method, make an interesting contribution.

There is certainly a need for a book of this kind, both as a laboratory manual and to provide critical appraisals of methods currently being used and developed. Although a broad selection of techniques is covered in the book, a number of other aspects of neurochemical methods would benefit from the same treatment, and I look forward to a future volume added to the *Practical Approach* series. □

J. de Belleruche is in the Department of Biochemistry, Charing Cross and Westminster Medical School, Fulham Palace Road, London W6 8RF, UK.

Horizons in molecular biology

Delegates to Nature's second conference in Japan were given ample opportunity to judge the powerful influence of molecular biology on a broad span of biological sciences.

WHAT is on the horizon in molecular biology? asked the speakers at the *Nature* conference in Tokyo last week (26–28 January). Their answers ranged from the complete mapping and sequencing of the human genome to genetically improved enzymes and plants, and from a deep understanding of how cell-surface receptors function to victory over AIDS.

Like it or not, it is certain that the techniques of molecular biology will continue to dominate biological research for many years to come. Views are mixed, however, as to how best to get back to 'real' biology. At one extreme there is the smash-and-grab approach, whereby molecular biology is used rapidly to gather information on, say the subtypes and structures of an enzyme, after which it is much more important to integrate that information back into a biological system than to continue on the seductive molecular biological track. At the other extreme there is the get-it-over-with approach that advocates sequencing, or making available for sequencing, the complete human genome as a priority. As Sydney Brenner (MRC, Cambridge) would have it, this is not a task fit for promising young men, who should be doing experimental biology, but is ideal for promising old men.

Brenner advocates that the first step should be to produce a complete and ordered library of the genome, in which the shelves are filled with consecutive, identified small pieces of the genome that can be loaned out on request. The pieces will be made at random but in a way that allows overlapping pieces to be recognized by computer analysis of the pattern of their restriction enzyme fragments separated by size on a gel. The technique has already been successfully applied to the genome of a nematode and several bacteria. Recently, an ordered library of about 70 per cent of the genome of a *Rhizobium* bacterium was produced in about a month. At that rate, it might take

ten people five years to near completion of a human genome library, but Brenner predicted that technical developments would reduce the time by an order of magnitude.

With his complementary approach, Charles Cantor (Columbia University) reckons it will take his group about a year to map 90 per cent of human chromosome 21, the smallest of them all. The technique, already used to map the entire genome of *Escherichia coli* and now capable of doing so in the course of a month, involves cutting the genome into relatively few but large fragments, separating them by pulsed field gel electrophoresis, and then identifying them and their neighbours by means of DNA probes. Along the way of mapping human chromosomes so far, the technique has both yielded valuable new polymorphic DNA markers and given a clear indication that house-keeping genes — those that encode everyday proteins — occur in clusters rather than at random.

For Akiyoshi Wada (University of Tokyo), the prime goal is large-scale fully automated DNA sequencing, the machinery for which is under development in Japanese government and industrial laboratories. Whereas a complete human sequence will take decades to complete and will include the 98 per cent of the genome that does not encode protein (but doubtless has many mysteries to reveal), Brenner argues that it is the coding 2 per cent, in the form of complementary DNAs (cDNAs), that should be sequenced first both for ease and because of its information density.

Just how many biological problems can stem even from limited cDNA cloning is illustrated by the flood of information on subtypes of protein kinase C, an enzyme that is involved in cell signal transduction and is the target for phorbol ester tumour promoters (Yasutomi Nishizuka, Kobe University). Cloning has identified at least

eight subtypes and there may be dozens to come. The question is, what different functions do they serve?

Although there are few answers, two clues have recently emerged from the Nishizuka laboratory. One is that the γ subtype of the enzyme is highly localized to the Purkinje cells of the cerebellum, where it may be involved in long term potentiation; the other is that γ , but not α or β subtypes are activated by quasi-physiological concentrations of arachidonic acid.

Equally challenging is the determina-



Robert H. Michell.

tion of the functions of the many inositol phosphates, which are largely derived along with diacylglycerol, the activator of protein kinase C, from the hydrolysis of phosphatidyl inositols. Whereas inositol triphosphates have received much attention, increasing attention is being given to, for example, a hexaphosphate that can sometimes be found in rather high concentrations, and to various isomers of the tetraphosphate that appear with different kinetics after vasopressin stimulation of mammary cells (Robert Michell, Birmingham University).

Protein structure

One long-standing goal of molecular biology is to understand fully the structural basis of enzyme action both to satisfy the intellect and with a view to improving enzymes for specific applications. The best situation is to start with an enzyme whose amino-acid sequence and three-dimensional structure are known, and for which there is a working model for how the enzyme functions, and a cloned cDNA that can be mutated in ways that will test the model's predictions. Such is the case with trypsin, whose activity, like that of all serine proteases, is believed to rely particularly on a serine together with a histidine and an aspartic acid in the active site. A series of single amino-acid substitutions, by means of site-directed mutagenesis, has confirmed the role of the three active site amino acids. More sur-



Tasuko Honjo (left). Nobuhiro Gō, Kurt Wuthrich and Haruo Ozeki.



Gustav Nossal.

prisingly, it has been shown that some of the amino acids lining the substrate-binding pocket of the enzyme are also intimately involved in catalytic function (William Rutter, University of California, San Francisco).

X-ray crystallography is still the only technique available for solving the structure of large proteins at an atomic resolution. The power of the technique in expert hands is illustrated by the continuing flow of information on the photosynthetic reaction centre of *Rhodospseudomonas viridis* (Hartmut Michel, MPI, Frankfurt). The methods developed to crystallize and analyse this membrane protein will sooner or later work for other membrane proteins, such as the many neurotransmitter receptors cloned in the laboratory of Shosaku Numa (Kyoto University).

At a lower resolution, image reconstruction from electron micrographs combined with labelling of specific sites can give valuable structural information. By such means, the actin binding-site on myosin has been identified by Takeyuki Wakabayashi (University of Tokyo) who is using undecagold-maleimide, which labels cysteine, to further his studies of actin and myosin. For assessing the dynamics of protein structures, computer simulation comes into play. Nobuhiro Gō (Kyoto University) showed some striking film of 'breathing' proteins, but it has to be remembered that the simulations cannot yet take account of water interactions.

Structures of proteins in solution are likely more accurately to represent the physiologically relevant structures than those of proteins in crystalline form. In that respect, two-dimensional nuclear magnetic resonance is becoming increasingly valuable, although it is still limited to rather small proteins (Kurt Wüthrich, ETH, Zurich). In general, the crystal and solution structures of globular proteins are rather similar, with any differences being on the surface of the proteins where, however, their importance may outweigh their scale. The technique has recently been extended to a DNA-binding fragment (of 76 amino acids) of a bacteriophage repressor. With synthetic DNA bound to the fragment, the spectrum becomes too complex for analysis. Introducing ^{15}N or ^{13}C amino acids into the repressor fragment makes it possible to deter-

mine some substructure that will assist in solving the complete structure.

A DNA-binding protein whose structure is eagerly awaited is TFIIA, the archetypal 'zinc-finger' activator of gene transcription. Along with two accessory proteins, TFIIA forms a complex that can bind to a regulatory region of both the oocyte-type and somatic-type 5S ribosomal RNA genes of *Xenopus*. The intriguing question is why it is largely the oocyte-type gene that is activated in the oocyte and the somatic-type gene later in development. Part of the answer must lie in the three-base difference in the TFIIA-binding site of the two genes and in histone H1-dependent mechanisms that can repress gene transcription, but there is also an unexplained mechanism by which the stability of different transcription complexes can vary with developmental stage (Donald Brown, Carnegie Institute of Washington).

Immunology

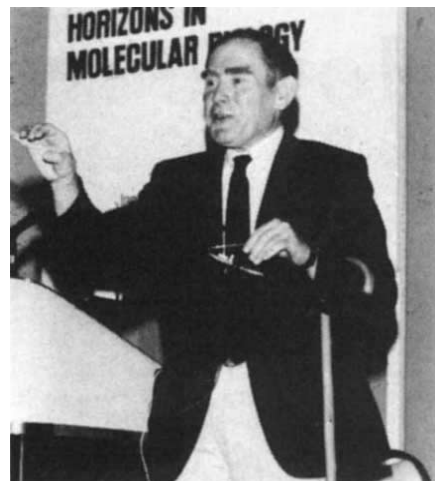
The immune system continues to be a meeting ground for various branches of biology, and for the application of molecular techniques. Gustav Nossal (Walter and Eliza Hall Institute, Melbourne) described experiments strongly supporting the notion of clonal anergy as the mechanism by which tolerance to 'self' is developed. The suppression of antibody production is greater for antibodies directed against cell-surface molecules than against internal proteins, which makes sense. The process of development from a pre-B cell into an antibody-producing



Robert Gallo.

lymphocyte continues to attract attention and a promising new system in which to study it was outlined by Tasuko Honjo (Kyoto University). He has developed pre-B cell lines that can be induced to mature *in vitro* by either bone marrow cells or a mixture of dendritic and T cells.

Immune system deficiencies are of more interest to Robert Gallo (National Cancer Institute), who sees some glimmer of hope for the development of vaccines against AIDS in evidence that individuals infected with one strain of virus seem to resist



Sydney Brenner.

infection by widely different strains. He also described evidence that the recently discovered human B-cell leukaemia virus may in fact have T4 cells as its main target and so may contribute to the development of AIDS after infection with human immunodeficiency virus. No new virus has emerged from the search he has led to account for Kaposi's sarcoma in AIDS patients. Instead, he has evidence of a growth factor produced by Kaposi sarcoma cell lines that will induce the sarcoma when injected into nude mice.

Plants

Molecular biology continues to have a powerful influence on both basic and applied plant science. Very promising field trials of plants that have been made resistant to specific pathogens or herbicides by transferring the relevant prokaryotic genes to their genomes were reported by Jozef St Schell (MPI, Köln), while Hirofumi Uchimiya (Tsukuba University) described the development of protoplast systems for the genetic engineering of two important Japanese crops — oranges and rice.

Schell also reported an analysis of the developmentally regulated gene for leghaemoglobin, which is expressed only in root nodules. Both an enhancer and a suppressor element have been identified, and an enhancer-binding protein that is present only in root nodules is being characterized. Aspects of the structure, organization and control of the genes of liverwort chloroplasts were discussed by Haruo Ozeki (Kyoto University); in many respects the system is more prokaryotic than eukaryotic. For example, a polycistronic messenger RNA has been identified. Several chloroplast proteins, such as ATPase, have a subunit structure, but in no case are the genes for the subunits all in the chloroplast genome; those in the nucleus are assumed to have migrated there at some time in the distant past. How the nuclear and chloroplast genes are coordinately regulated remains a mystery.

Peter Newmark

What the universities have done

Mark Richmond

The British government says that the universities must adapt to the nation's changing needs. They are already doing so, and implementation of the Education Reform Bill will only hamper the process.

IT is common ground between the universities and the government that Britain's future depends upon the existence of a highly educated and well-trained workforce. Both of us recognize that civilized values rest on adequate wealth creation, and that that is dependent on a strong industrial base at the forefront of technological development.

The universities have a crucial role in underpinning wealth creation. They do it not only by the education of young people to play an effective part in the world of work, but through being the places where much of the fundamental research is carried out from which technological advance springs.

So the universities are at the heart of the matter. But things change, and nothing is ever entirely or lastingly 'right'. There is certainly a case for arguing that in the post-Robbins era, during the 1960s and 1970s, the universities were too distant from the immediate needs of the country, and that higher education was concentrated on too narrow a section of society. The universities welcome the government's commitment to higher standards in education, to increased access and to the stimulation of entrepreneurial endeavour. We are keen to play a larger part in educating people of all ages and to support the research and development that will underpin the modernization and growth of British industry.

To a degree, all this is threatened by proposals put forward in the Education Reform Bill, now before parliament (see *Nature* 331, 287 and 291; 1988). In asserting this, I stress we are not seeking to protect privilege or to secure special status for the universities or their staffs. Rather, we are arguing vehemently for three fundamental freedoms which we believe to be of lasting value to Britain — the freedom to research in subjects of as yet unrecognized importance, the freedom to question received wisdom, and the freedom to be protected from direct and narrow political interference by the government of the day. In fighting for these

freedoms, we are not arguing against the need for change. We are, however, arguing that the provisions of the Education Reform Bill are not necessary to produce what is needed.

My purpose in this article is to demonstrate that the universities are adapting, that there has been a response to the government's lead — just as in the 1960s the

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Building bridges — the 'launch' of Salford University's Business Park in October of last year. The intention is to attract small, high-technology companies who will benefit from close co-operation with the university.

universities responded to different leads from a different government.

First, the money. In 1980–1981 the universities' recurrent grant was cut substantially (Fig. 1 — see panel overleaf). Our protests are well known and the government, to some extent, has relented. The last Autumn Statement brought good news — albeit against a grisly background — and next year we will be virtually back in real terms to where we were in 1980. Real growth is planned in the next two years.

Faced with this trough, the universities' response has been interesting. Since 1980 the number of teaching staff has fallen by 12 per cent (Fig. 2), which has been achieved without compulsory redundancies. Initially, as the cuts bit, the number of full-time students also fell. Indeed we were instructed to cut numbers by the University Grants Committee, and those who did not were fined in the following year. After two or three years this policy was reversed, numbers began to go up

again and reached an all-time record of over 301,000 by the academic year 1986–1987 (Fig. 3). As shown in Fig. 4, the number of part-time students has gone up more sharply (here it should be remembered that the Committee of Vice-Chancellors and Principals includes among its membership the Open University, the biggest part-time higher education institution in Britain). Short courses, often run for industry and commerce, are increasing too. They have become an increasingly important source of income (Fig. 5).

What all this adds up to is that fewer staff are doing more teaching. But there is also more research going on, so the universities are not doing the teaching at the expense of research (Fig. 6). Total research income is up by 85.7 per cent, and (Fig. 7) within that total earnings from industry and commerce have gone up particularly sharply, much faster in fact than income from the Research Councils.

Nor is that all. Technology transfer activities — consultancy and applied research — which hardly existed before 1980, have burgeoned. There are, for example, about 100 university companies providing these sort of services, and on top of all that well over 500 spin-off companies have been set up by British universities in the past few years.

More teaching, more research, more technology transfer: it all amounts to a major gain in productivity, and we are confident that there has been no collapse in quality. This gain has been helped by streamlining of financial management, following the recommendations of the Jarratt Report of 1985. This report was commissioned by the universities themselves. Standardized accountancy practices are being implemented, and we are devising new forms of management statistics which will be used by institutions to monitor their own efficiency and effectiveness.

On the academic side, procedures for monitoring standards have been tightened up and will be tightened even more. Courses are being revised and updated; subject departments are being reorganized to meet changing patterns of demand.

Another graph indicates how necessity has been the mother of invention, how by changing the product to suit the market the universities have managed to grow.

• Erratum: In the editorial on this subject in last week's *Nature*, a 'not' strayed into the last sentence of the first column. The passage should have read "... there is general agreement that the Education Reform Bill is, so far as the universities are concerned, a direct threat to their independence, which must be resisted."

Figure 8 shows what has happened to the number of overseas students in British universities since subsidies for such students were withdrawn in 1979. Initially there was a sharp fall. Quickly new courses and new marketing drives were introduced, and numbers are now at an all-time high.

Universities have long provided expert advice and service — forensic scientists, lawyers and economic analysts, for example — and they have worked closely for many years with the main science-based companies such as ICI, BP, Shell and those in the pharmaceutical industry. What is new is that universities are now actively involved with many hundreds of smaller companies as well: old companies which are modernizing or new companies starting up in high-technology areas as the economy picks up. The relationship between these new customers and the universities is very exciting — here is a highly cost-effective way of generating healthy science- and technology-based industrial growth.

This entrepreneurial activity, this ability to change, this ability to generate private income, has saved a number of universities from bankruptcy and all of them from stagnation. A further consequence is that the university system is now much more vital and diverse. From the perspective of the government these must have been major changes in the right direction — closer to industry and towards less dependency on the Exchequer. The proportion of the universities' income coming from exchequer grants is now down to 57 per cent compared to 77 per cent a decade ago (Figs 9 and 10).

It is the centre of the universities' case that we have been able to do all this because we have been masters of our own houses. We agree with the Prime Minister that the government that governs best is the government that governs least. The effectiveness of our response over the past few years derives in great measure from our freedoms. Those freedoms are now threatened. The Education Reform Bill includes among its clauses a provision that the new council to be set up to channel government funds to universities shall comply with *any* directions given to it by the Secretary of State for Education and Science. The council in its turn is to have the power to require any information from universities; to set any terms and conditions; and to require repayment with interest of *all* money if the conditions are not met. That will not help the good and enterprising managers to develop universities to meet Britain's future needs. □

Sir Mark Richmond is Vice-Chancellor of the Victoria University of Manchester and Chairman of the Committee of Vice-Chancellors and Principals. This article is an edited version of part of his address given on 19 January at a seminar on the Education Reform Bill.

Change in the 1980s: funds, staff, students

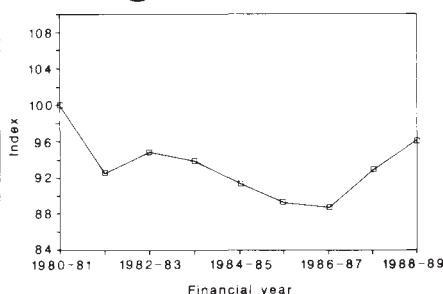


Fig. 1 Universities' recurrent grant, 1980-81 to 1988-89, adjusted for funding changes.*

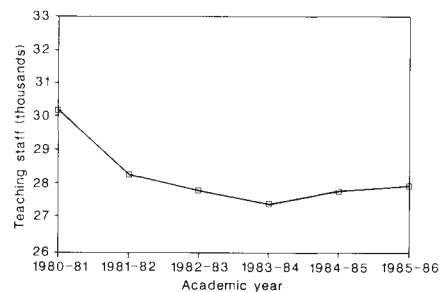


Fig. 2 Numbers of non-clinical teaching staff, 1980-81 to 1985-86.

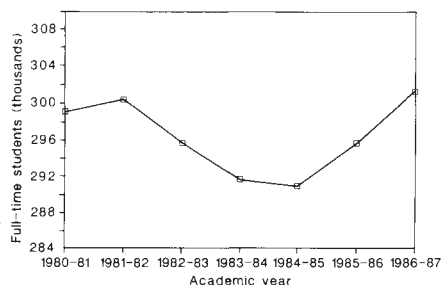


Fig. 3 Numbers of full-time students, 1980-81 to 1986-87.

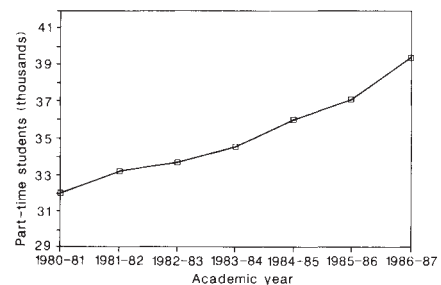


Fig. 4 Numbers of part-time students, 1980-81 to 1986-87.

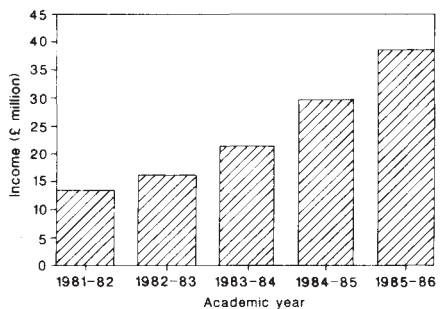


Fig. 5 Income from short courses, 1981-82 to 1985-86.

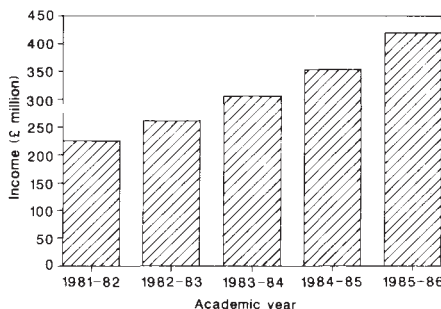


Fig. 6 Income from research grants and contracts, 1981-82 to 1985-86.

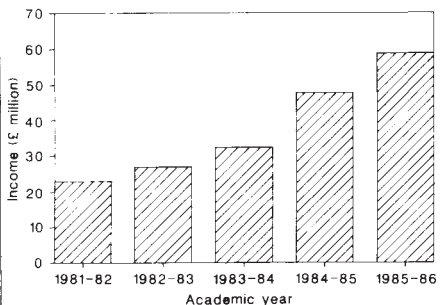


Fig. 7 Income from industry and commerce, 1981-82 to 1985-86.

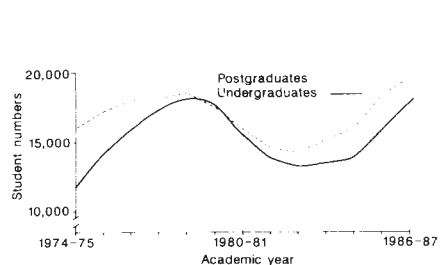


Fig. 8 Numbers of overseas students, 1974-75 to 1986-87.

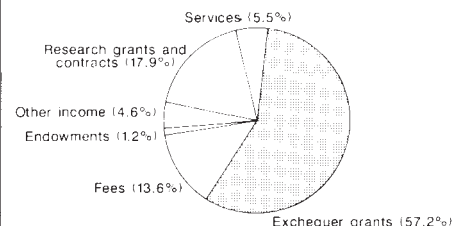


Fig. 9 Universities' recurrent income 1985-86.

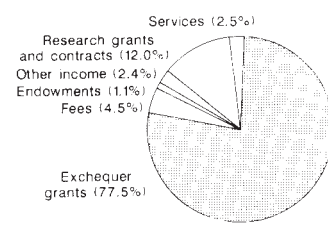


Fig. 10 Universities' recurrent income 1974-75.

*Source: UGC. Source for the other diagrams is *University Statistics*, Vols 1 and 3 (*University Statistical Record*).

'Domestic' origin of opaque assemblages in refractory inclusions in meteorites

Joel D. Blum, G. J. Wasserburg, Ian D. Hutcheon, John R. Beckett & Edward M. Stolper

Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA

Experimental studies indicate that opaque assemblages rich in refractory siderophile elements were formed within host calcium- and aluminium-rich inclusions (CAIs) by exsolution, oxidation and sulphidization of homogeneous alloys, rather than by aggregation of materials in the solar nebula before the formation of CAIs. These opaque assemblages are thus not the oldest known solid materials, as was once thought, and they do not constrain processes in the early solar nebula before CAI formation. Instead, the assemblages record the changing oxygen fugacity experienced by CAIs during slow cooling in nebular and/or planetary environments.

CALCIUM- and aluminium-rich inclusions (CAIs) in C2 and C3 chondritic meteorites are generally believed to be representative of the earliest solids formed in the primitive solar nebula^{1,2}. They are enriched relative to cosmic proportions in the elements that are most refractory in a gas of solar composition. Reflecting this trend is an enrichment ($\sim 20\times$ CI chondrite) in the refractory siderophile elements (Ru, Os, Re, Pt, Ir, W and Mo)³⁻⁵, which are concentrated in multi-phase opaque assemblages⁶⁻⁸. The bulk chemistry of these assemblages has strong affinities with the calculated earliest condensates from a hot portion of the solar nebula^{5,6,13}. Extensive investigations have shown that many of the assemblages are composed of discrete micrometre-sized refractory siderophile metal nuggets surrounded by Ni-Fe metal, V-rich magnetite and Fe-Ni sulphides with associated molybdenite, molybdate, tungstate and phosphate⁹⁻¹³. The opaque assemblages are typically spheroidal and 1–1,000 μm in diameter. They occur as inclusions in all of the major silicate phases of CAIs, are sometimes found in embayments in spinel, and are often rimmed by a thin ($\leq 20\ \mu\text{m}$) rind of V-rich fassaite^{9,11,12} (see Fig. 1). In addition to the opaque assemblages described above, there exist other small (0.5–3.0 μm) refractory siderophile metal nuggets not associated with Ni-Fe metal, magnetite or sulphide^{7,10,13,37}. The net refractory siderophile element content of CAIs is controlled by the combination of these smaller objects and the opaque assemblages described here. We will focus here on the origin of the opaque assemblages.

Owing in part to the presumption that refractory siderophile metal nuggets formed either as condensates in expanding supernova envelopes (outside the Solar System) or as the earliest high-temperature nebular condensates, early workers postulated that the opaque assemblages were exotic and formed before CAIs⁷⁻⁹. As a result, the opaque assemblages are often referred to by the genetic name "Fremdlinge"^{8,9}, which is German for strangers or foreigners. We prefer to refer to the objects, previously called "Fremdlinge", by the simple descriptive name "opaque assemblages".

Recent studies of the chemical compositions, mineralogy and textural relationships of opaque assemblages have led most workers to concur that opaque assemblages formed before the host CAIs, and to agree on the following scenario for their origin and evolution¹¹⁻¹³: (1) condensation of Ru,Os-rich nuggets at very high temperature ($\sim 1,500$ – $2,000\ \text{K}$) in a supernova envelope or in a hot part of the solar nebula; (2) condensation of Ni-Fe metal followed either by oxidation of Fe and reaction with V in a nebular gas to produce V-rich magnetite or by direct condensation of V-rich magnetite; (3) aggregation of Ru,Os-rich nuggets, Ni-Fe metal, V-rich magnetite and associated phases in the nebula at low temperature ($\sim 870\ \text{K}$) to form opaque

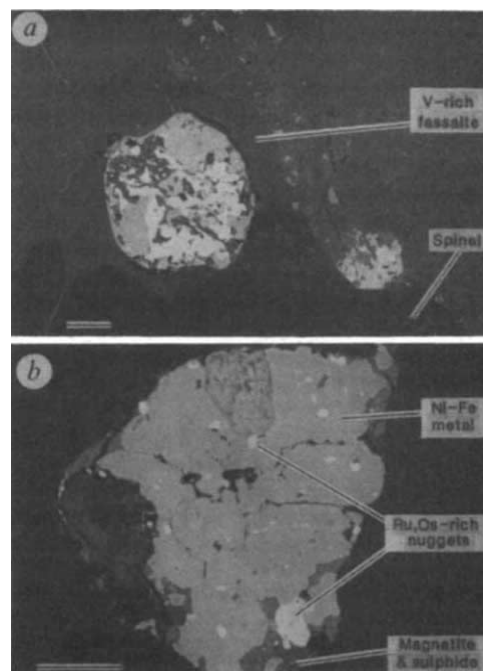


Fig. 1 Backscattered-electron images of opaque assemblages from CAIs. *a*, Two opaque assemblages (round bright objects) which occur in embayments in a spinel crystal in Allende sample USNM 5241. The surrounding silicate is melilite and the rim around the larger opaque assemblage is V-rich fassaite. *b*, An opaque assemblage in Allende JIM analysed in this study, which includes Ru,Os-rich nuggets, Ni-Fe metal, V-rich magnetite and sulphides. Scale bars, 10 μm .

assemblages; and (4) mixture of opaque assemblages with proto-CAI material before (or during) a brief high-temperature ($\sim 1,700\ \text{K}$) event, during which the silicate portion of the host CAI melted or partially melted, followed by rapid cooling. The high-temperature CAI melting event was postulated to have been so brief, and cooling so rapid, that the delicate textures and mineral intergrowths in opaque assemblages were preserved as they existed before CAI melting. Thus, opaque assemblages have been inferred to record conditions in the solar nebula before CAI crystallization, providing important constraints on the temperature, composition and degree of mixing in the early nebula^{9,11-13,15}.

To test experimentally models for the formation of opaque

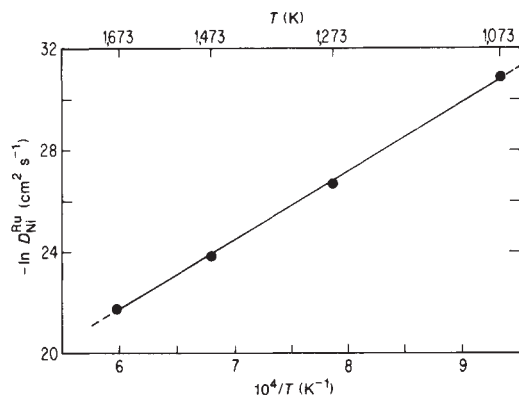


Fig. 2 Summary of the results of Ru-Ni thin-film diffusion experiments at four different temperatures, yielding a reliable estimate of the diffusion coefficient (D_{Ni}^{Ru}) as a function of temperature (T), based on the fitted line. The equation of the line is $D_{Ni}^{Ru} = 0.005 \exp(-2.3 \times 10^{12}/RT)$; the activation energy thus determined is $(2.3 \pm 0.1) \times 10^{12}$ erg mol $^{-1}$, and R is the gas constant in c.g.s. units.

assemblages, we have determined the diffusion coefficient for Ru in Ni (D_{Ni}^{Ru}) as a function of temperature, studied phase equilibria in the system Ni-Fe-Ru-O, and studied the partitioning of minor amounts of Pt, Ir and V between coexisting phases in systems dominated by Ni, Fe, Ru and O. Based on our experiments, we propose a new model for the origin of opaque assemblages that is in sharp contrast with the prevailing scenario¹¹⁻¹³. We suggest that opaque assemblages originated as homogeneous metallic droplets in CAI silicate liquids and that, after crystallization and subsequent cooling to a temperature of ≈ 870 K, these once homogeneous alloys exsolved into immiscible metallic phases as they were partially oxidized and sulphidized to form the multi-phase opaque assemblages now observed in CAIs. The possibility that some refractory siderophile metal nuggets in opaque assemblages might have formed by exsolution was suggested previously^{6,10,14}, but this view was subsequently rejected by most workers in favour of the formation of these nuggets as primary condensates^{9,11-13,15}, and has therefore received little serious attention until now. The implications of our model are that opaque assemblages do not constrain the high-temperature thermal histories of CAIs or conditions in the solar nebula before CAI melting, as suggested previously. Instead, they provide important constraints on the low-temperature thermal histories of CAIs and on the changing oxygen fugacity (f_{O_2}) in cooling nebular and/or planetary environments.

Experiments and analyses

Thin-film diffusion experiments were conducted by electroplating a 0.3- μ m film of Ru onto a slab of polycrystalline Ni. Pieces of this material were annealed at 1,673, 1,473, 1,273 and 1,073 K in evacuated silica tubes with graphite for times ranging from 0.3 to 137 h. The samples were sectioned and Ru concentration ([Ru]) profiles measured at 1- μ m intervals perpendicular to the interface, using a JEOL 733 electron microprobe following procedures given in ref. 11. The value of D_{Ni}^{Ru} was determined at each temperature (T) from the approximately linear slope of $-\ln[Ru]$ versus the square of the distance, following the thin-film approximation¹⁶. The values of D_{Ni}^{Ru} determined in this way are shown plotted against $1/T$ in Fig. 2 and yield the expression for $D_{Ni}^{Ru}(T)$ given in the Fig. 2 legend.

To study Ni-Fe-Ru-O phase equilibria and Pt and Ir partitioning, we first produced a powdered metal mixture with the composition $Ni_{58}Fe_{40}Pt_2$ (in atomic proportions). To aliquots of this mixture, we added 10% Ru, 20% Ru, 10% Ru + 5% Ir, and 20% Ru + 5% Ir. The powders were melted and homogenized. Pieces of this material were then sealed in evacuated silica

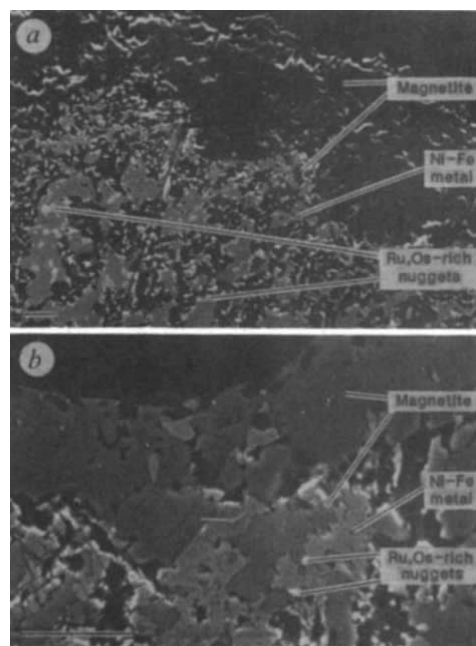


Fig. 3 Secondary-electron images comparing a typical experimental run product (a) and a portion of the opaque assemblage "Willy"¹¹ from Allende USNM 5241 (b). Note the textural and chemical similarities, which include: (1) a magnetite rim and intergrowths; (2) Ni-Fe metal with dissolved Pt and Ir; and (3) Ru $\pm$ (Os, Ir)-rich nuggets. Scale bars, 10 μ m.

tubes with a Ni-NiO buffer and annealed at 1,273 and 1,073 K for 1 and 9 days respectively. Reaction of oxygen with each of the synthetic alloys produced an outer rind of magnetite and a zone of magnetite (or in some instances wüstite) intergrown with Ni-Fe-Pt-Ru($\pm$ Ir) alloys just inside the rind (Fig. 3). The removal of Fe from metal to form magnetite produced metal enriched in Ni, Pt, Ru, $\pm$ Ir near regions where magnetite had formed, because these metals have very low solubilities in magnetite. A compositional gradient was thus established, and where the metal composition became sufficiently enriched in Ni and Ru, a Ru-rich metallic phase (nuggets) exsolved. Coexisting Ru-rich nuggets and Ni-Fe-rich metals in synthetic samples were analysed using a JEOL JSM 35 analytical scanning electron microscope (ASEM) following procedures given in ref. 11. Only exsolved nuggets larger than the ~ 0.8 - μ m X-ray interaction volume were analysed, to minimize interference from surrounding phases. Pt was concentrated in the Ni-Fe metal, whereas Ir partitioned into both the Ni-Fe metal and the exsolved Ru-rich nuggets. The positions of phase boundaries and tie-lines at 1,273 and 1,073 K in the Ni-Fe-Ru phase diagram were determined by analysing the compositions of coexisting Ru-rich nuggets and Ni-Fe metals and assuming local equilibrium on a 2-3- μ m scale. The experimentally determined 1,273- and 1,073-K phase diagrams, and the 873-K diagram inferred from the three binary systems, are shown projected from the Pt apex in Fig. 4.

To study V partitioning between Ni-Fe metal and magnetite, we dissolved V into Ni-Fe under f_{O_2} conditions one log unit more oxidizing than the solar gas curve (Fig. 5) to form a $Ni_{38}Fe_{61}V_1$ alloy. We annealed this alloy under the same experimental conditions as the Ni-Fe-Pt-Ru($\pm$ Ir) alloys. Essentially all of the V initially in the metal alloy partitioned into the magnetite rim and intergrowths, resulting in V-free ($<0.1\%$) Ni-Fe metal and V-rich magnetite.

To compare the experimental run products with naturally occurring opaque assemblages in CAIs, three opaque assemblages from Allende sample JIM¹⁷ (Fig. 1b), Allende USNM

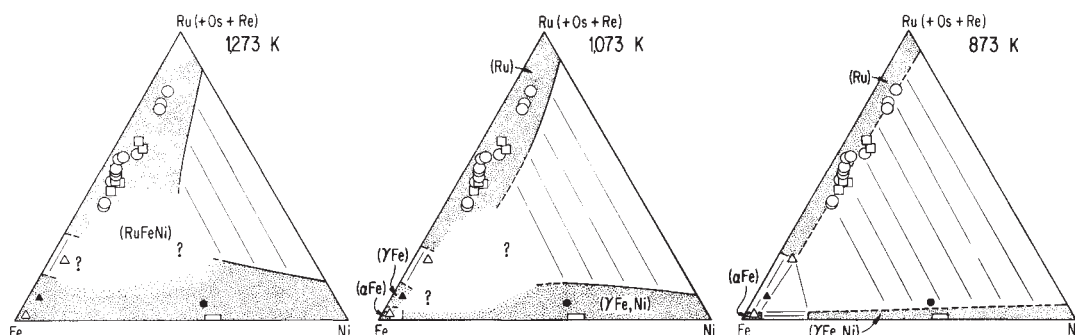


Fig. 4 Isothermal sections for the sub-solidus phase diagram in the Ni-Fe-Ru system (projected from Pt). Experimentally determined phase boundaries (including data from the three binary systems^{19,39,40}) are shown as dark solid lines; dark dashed lines are phase boundaries extrapolated into compositional regions not yet investigated. Phase diagrams are incomplete (denoted by “?”) in the Fe-rich region at 1,273 and 1,073 K, where several constructions are possible. Single-phase regions are shaded. Tie-lines are shown in two-phase regions as light solid lines; they were determined experimentally where they connect solid phase boundaries and were estimated where they connect dashed phase boundaries. For comparison, analyses of metal phases (open symbols) from three opaque assemblages in CAIs (Allende JIM (○), Allende USNM 5241 “Willy”¹¹ (□) and Leoville UNM 575 (△)) are shown. The rectangular region near the Ni-Fe edge of the triangle encloses 42 analyses of Ni-Fe metal from both Allende opaque assemblages. The approximate bulk metal composition of the two Allende opaque assemblages (●) and the Leoville opaque assemblage (▲) are also shown.

5241 (“Willy”)¹¹ (Fig. 3b) and Leoville UNM 575¹⁸ were studied. Each comprises mostly Ni-Fe metal, V-rich magnetite and Ru,Os-rich nuggets. The opaque assemblage from Leoville UNM 575 has a $2 \times 10\text{-}\mu\text{m}$ Ru,Os-rich lamella rather than discrete nuggets. Most Ru,Os-rich nuggets in the two Allende opaque assemblages are $0.1\text{--}1\text{ }\mu\text{m}$ in diameter, allowing no more than a single analysis of the largest nuggets. One unusually large $3 \times 5\text{-}\mu\text{m}$ Ru,Os-rich nugget was found in the opaque assemblage from Allende JIM and four non-overlapping analyses were made on this nugget. Compositions of Ru,Os-rich nuggets and the lamella, and adjacent Ni-Fe metal, were measured by ASEM and bulk opaque-assemblage metal compositions were calculated from point-counts. Pt occurs only in the Ni-Fe metal; Ru,Os and Re occur almost entirely in the Ru,Os-rich nuggets; and Ir occurs in both Ni-Fe metal and Ru,Os-rich nuggets. We have represented the sum of the three chemically similar hexagonal-close-packed noble metals (Ru + Os + Re) as the Ru apex in Fig. 4, so that the compositions of Ru,Os-rich nuggets from opaque assemblages can be compared with the Ni-Fe-Ru phase diagrams. The compositions of Ru,Os-rich nuggets from the two Allende opaque assemblages (open circles and squares) spread along a linear trend of constant Ni concentration with variable proportions of Ru and Fe, whereas the Ni-Fe metal (open rectangle) has a fairly constant $\text{Ni}_{67}\text{Fe}_{33}$ composition. The Ru,Os-rich nugget compositions are lower in Ni than the composition that would be predicted to be in equilibrium with Ni-Fe metal at 1,073 K based on our experimentally determined phase boundaries in the 1,073-K diagram (Fig. 4). However, inspection of the 1,273- and 1,073-K diagrams shows that equilibrium nugget compositions contain progressively less Ni with decreasing temperature. Based on this trend and the inferred 873-K diagram, we suggest that Ru,Os-rich nuggets equilibrated with Ni-Fe metal at $\approx 873\text{ K}$. The nearly constant Ni-Fe metal composition could not have been in equilibrium with most of the Ru,Os-rich nugget compositions. We attribute this to homogenization of Ni-Fe metal after equilibration with Ru,Os-rich nuggets, as discussed below.

A good test of the applicability of the inferred phase diagram is the composition of the two coexisting metals in the opaque assemblage from the Leoville CAI, because the bulk metal composition is extremely Ni-poor and lies close to the previously determined Fe-Ru binary system¹⁹. If the two metal phases were once in equilibrium, their compositions should plot at the ends of a tie-line crossing the bulk metal composition at a particular temperature. As can be seen in Fig. 4, the Leoville opaque

assemblage metal compositions (open triangles) agree closely with equilibrium at 873 K.

Diffusion constraints

Arrhenius and Raub²⁰ first pointed out that the existence of sharp contacts between Ru,Os-rich or Pt-rich nuggets and Ni-Fe metal in opaque assemblages provides an important constraint on the cooling history of opaque assemblages and host CAIs. This argument was largely ignored, however, because $D_{\text{Ni}}^{\text{Ru}}$ was not known²⁰ and although $D_{\text{Ni}}^{\text{Pt}}$ was known, the existence of primary contacts between Pt-rich nuggets and Ni-Fe metal has never been confirmed¹¹. The expression for $D_{\text{Ni}}^{\text{Ru}}(T)$ determined here now permits a calculation of the initial cooling rate (r_0) required to preserve contacts between Ru,Os-rich nuggets and Ni-Fe metal. Using the maximum Ru concentrations in Ni-Fe metal measured 0.5 and $1.0\text{ }\mu\text{m}$ from Ru,Os-rich nuggets, and the equation for diffusion from a $1\text{-}\mu\text{m}$ -diameter sphere¹⁶, we first calculated the value for the time integral of $D_{\text{Ni}}^{\text{Ru}}(T)$ (or $\tau(\infty)$) that would produce the observed gradients²¹. Equating this to the approximation for $\tau(\infty)$ given by Kaiser and Wasserburg²² ($\tau(\infty) \approx D(T_0)RT_0^2/Er_0$, where R is the gas constant and E is the activation energy), and assuming that the maximum temperature of melting (T_0) was $\sim 1,700\text{ K}$, we calculate $r_0 \geq 10^5\text{ K h}^{-1}$. Although this r_0 is possible for CAIs in a radiative environment²³, it is unreasonably rapid in light of experimental studies of textures and phase chemistries of the silicate portion of CAIs, which imply much slower cooling ($r_0 \approx 0.5\text{--}20\text{ K h}^{-1}$; refs 23–27). On the other hand, we calculate from $D_{\text{Ni}}^{\text{Ru}}(T)$ that the diffusion of Ru in Ni is sufficiently rapid for exsolution of typical $1\text{-}\mu\text{m}$ Ru,Os-rich nuggets (at $10\text{-}\mu\text{m}$ spacings) to occur on a reasonable timescale of about 10 days or 10 years, depending on whether exsolution began at 1,073 or 873 K.

Origin of opaque assemblages

In our phase equilibrium and partitioning experiments, we began with homogeneous Ni-Fe-V alloys and produced V-rich magnetite as rims and intergrowths within V-poor Ni-Fe metal. We also began with Ni-Fe-Pt-Ru ($\pm$ Ir) alloys and produced magnetite as rims and intergrowths within Ni-Fe metal enriched in Ni and Ru. Ru-rich nuggets were produced by exsolution, and we observed that most Pt remained in the Ni-Fe phase, whereas Ir partitioned into both phases. The textures and mineral chemistries that we produced in the laboratory show a remarkable resemblance to natural opaque assemblages (Fig. 3), and lead us to suggest a new model for their origin. Following previous

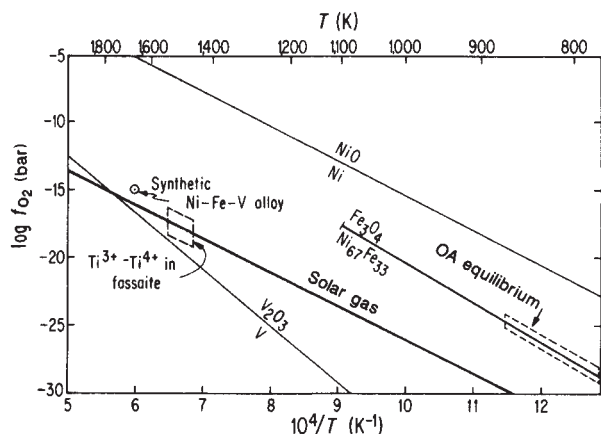


Fig. 5 Plot of $\log f_{\text{O}_2}$ versus $1/T$ with relevant equilibrium curves. The conditions of equilibrium for opaque assemblages (OA) are from ref. 11 and this work; the solar gas curve is at $P = 10^{-3}$ bar⁴¹; the $\text{V}/\text{V}_2\text{O}_3$ curve is calculated from thermodynamic data⁴²; the $\text{Ni}_{67}\text{Fe}_{33}/\text{Fe}_3\text{O}_4$ curve is from ref. 30; the fassaite crystallization conditions are based on $\text{Ti}^{3+} - \text{Ti}^{4+}$ (ref. 36); the open circle gives the experimental conditions for synthesis of the Ni-Fe-V alloy used as starting material in our V partitioning experiments.

workers, we assume that proto-CAIs are refractory, early nebular condensates or evaporative residues. We propose that when the silicate portions of CAIs melted, refractory siderophile elements which had previously condensed as refractory metal alloys^{5,6,13} partitioned into homogeneous metallic droplets: At the temperature of CAI melting ($\sim 1,700$ K), Ni-Fe metal would occur as molten droplets, consistent with the spheroidal shapes of many opaque assemblages. At this temperature, the only solid phase in most CAIs would be spinel²⁴. Subsequent crystallization of spinel could result in the occurrence of opaque assemblages in embayments in spinel crystals (Fig. 1a). As CAIs cooled from $\sim 1,700$ to $\sim 1,500$ K, the metallic droplets crystallized as homogeneous alloys and silicate minerals crystallized around them. At the still lower sub-solidus temperature of ≤ 870 K, the homogeneous alloys exsolved into immiscible metallic phases (such as Ru,Os-rich nuggets and Ni-Fe metal) as they were partially oxidized and sulphidized to form magnetite and associated phases, completing the formation of opaque assemblages.

An important requirement of this model is that there must be a substantial change in f_{O_2} during the formation of the opaque assemblages within the host CAIs. This change is required to explain several basic features of the opaque assemblages, including the oxidation of Ni-Fe metal to magnetite and the initial partitioning of substantial amounts of V into the metal phase at high temperature, followed by subsequent oxidation and diffusion into magnetite. In the remaining discussion, we address the origin of some of the detailed characteristics of opaque assemblages in the context of this model.

Ru,Os-rich nuggets

Ru,Os-rich nuggets from individual opaque assemblages show a large variability in $(\text{Ru} + \text{Os} + \text{Re})/\text{Fe}$ ratios. The compositions spread along the inferred 873-K phase boundary in the Ni-Fe-Ru diagram (Fig. 4). This behaviour is consistent with the exsolution of the Ru,Os-rich nuggets in stages during the removal of Fe from Ni-Fe metal to form magnetite and Fe sulphide, so that a variety of nugget compositions formed as the bulk metal composition became more Ni-rich. If the Ru,Os-rich nuggets were the product of simple equilibrium exsolution, there would be a single nugget composition falling on a common tie-line crossing the bulk metal composition. Only one Ru,Os-rich nugget was large enough to permit multiple analyses; analyses of four separate areas indicate that this nugget is heterogeneous in composition²¹. It is possible that all nuggets are somewhat zoned, and we expect that the edges of each nugget have a constant composition in equilibrium with adjacent $\text{Ni}_{67}\text{Fe}_{33}$.

Ni-Fe metal

The Ni-Fe metal compositions of opaque assemblages in most Allende CAIs are fairly constant ($\text{Ni}_{67}\text{Fe}_{33}$), in contrast to the variable compositions of Ru,Os-rich nuggets. We consider $\text{Ni}_{67}\text{Fe}_{33}$ to reflect the Ni-Fe metal composition in equilibrium

with magnetite at the f_{O_2} of the local CAI environment, and thus to be the endpoint of the oxidation process. We suggest that the Ni-Fe metal homogenized during oxidation and exsolution, whereas the isolated Ru,Os-rich nuggets were kinetically inhibited from re-equilibrating with each other. This is consistent with the lower value of $D_{\text{Ni}}^{\text{Ru}}$ at 873 K determined here ($\sim 2 \times 10^{-16} \text{ cm}^2 \text{ s}^{-1}$), compared to the value of $D_{\text{Ni-Fe}}^{\text{Ni}}$ at 873 K determined by Hirano *et al.*²⁸ ($\sim 3 \times 10^{-15} \text{ cm}^2 \text{ s}^{-1}$). We note that $\text{Ni}_{67}\text{Fe}_{33}$ alloy also coexists with magnetite and sulphide in opaque spheroids in olivine-rich chondrules from Allende^{29,30}, suggesting that both chondrules and CAIs in this meteorite were exposed to the same ambient f_{O_2} and that the opaque assemblages in the chondrules also represent an oxidation endpoint.

There is a substantial discrepancy between the predictions of equilibrium condensation from a cooling gas of solar composition and the composition of Ni-Fe metal observed in Allende, as models predict a maximum Ni content of about $\text{Ni}_{20}\text{Fe}_{80}$, whereas $\text{Ni}_{67}\text{Fe}_{33}$ is observed^{31,32}. This discrepancy can be partly accounted for by the enrichment of opaque-assemblage metal in Ni caused by the removal of Fe during oxidation. But adding the Fe contribution from magnetite and Fe sulphide in Allende opaque assemblages back into the bulk metal composition only brings the composition back to a minimum Ni content of about $\text{Ni}_{33}\text{Fe}_{67}$, which is still somewhat more Ni-rich than the $\text{Ni}_{20}\text{Fe}_{80}$ condensate composition. If metal condensation occurred under conditions more oxidizing than a solar gas, condensation of more Ni-rich metal is theoretically possible¹³.

V-rich phases

The V-rich fassaite rims that frequently surround opaque assemblages can also be understood in the context of our model. We suggest that metallic V was present in molten Ni-Fe alloy droplets at high temperature, in contrast to previous models that assume that V was always present as V_2O_3 , which is the oxidized form of V known to occur in silicate and oxide phases^{11,12}. In our view, metallic V condensed from the solar nebula either as a metal, or as an oxide which was subsequently reduced during the melting of CAIs. Consider a plot of $\log f_{\text{O}_2}$ versus $1/T$ that includes both the solar gas curve (at $P = 10^{-3}$ bar) and the equilibrium curve for the reaction $2\text{V} + \frac{3}{2}\text{O}_2 = \text{V}_2\text{O}_3$ (Fig. 5). The two curves cross at $T \approx 1,770$ K and $\log f_{\text{O}_2} \approx -15$, so that for $T \geq 1,770$ K metallic V is stable at the f_{O_2} of a solar gas. A detailed calculation of the $\log f_{\text{O}_2}$ versus $1/T$ curve along which an opaque-assemblage alloy with metallic V would be in equilibrium with a CAI silicate liquid containing oxidized V, indicates that this curve would be displaced somewhat to the right of the $\text{V}/\text{V}_2\text{O}_3$ curve. The shift would lower the temperature at which V would be stable in a Ni-Fe metal droplet immersed in CAI silicate liquid at the f_{O_2} of a solar gas to $\sim 1,670$ K. This calculation uses available estimates of V activities in metal³³ and silicate³⁴, V concentrations in opaque assemblages and bulk

CAIs¹¹, and an estimate of the fraction of atoms in CAIs once contained in opaque assemblages¹¹. As CAIs cooled along the solar gas curve, V in refractory siderophile alloys would be expected to oxidize progressively and diffuse out of the metal because of the difference in the slopes of the solar gas and V/V_2O_3 curves (Fig. 5). The oxidized V could have diffused into surrounding fassaite, resulting in the V-rich fassaite rims. The oxidized V in the silicates may have been present as VO rather than V_2O_3 at high temperatures. But as the V/VO curve is only ~ 1 log unit more reducing than the V/V_2O_3 curve, the change in speciation would not significantly affect the above argument.

All of these phases in opaque assemblages cannot be produced simply by cooling of a solar gas, because the $Ni_{67}Fe_{33}$ metal and V-rich magnetite must have equilibrated at an f_{O_2} that is ~ 6 log units more oxidizing than the solar gas curve at a temperature of ~ 870 K (Fig. 5). The deviation from the solar gas curve that led to the formation of V-rich magnetite at ~ 870 K could also have influenced the oxidation of V at intermediate temperatures. Nevertheless, in the context of our model for the origin and evolution of opaque assemblages, the f_{O_2} of the early high-temperature solar nebula must have been at least as reducing as the solar gas curve during CAI melting because of the necessity for V to occur in the alloy. This f_{O_2} is consistent with estimates of the f_{O_2} during CAI crystallization based on $Ti^{3+} - Ti^{4+}$ in fassaite^{35,36} (Fig. 5). In these reducing conditions, the siderophile elements Mo, W and P would also have been in a reduced state in the opaque-assemblage alloy. By the time CAIs cooled to ~ 870 K, however, the f_{O_2} must have deviated from the solar gas curve and increased dramatically to form the V-rich magnetite and possibly the molybdate, tungstate and phosphate found in opaque assemblages.

Other phases

In addition to the coexistence of Ru,Os-rich nuggets and Ni-Fe metal, there are other intergrown phases in opaque assemblages that indicate equilibrium at ~ 870 K (for example, $Ni_{67}Fe_{33} - Fe_3O_4 - FeNiS$, $Fe - FeS - MoS_2$ and $Fe - Fe_3O_4 - FeAl_2O_4$)¹¹. Although we have not explored diffusion within these phases experimentally, we consider it unlikely that intergrowths of these phases could have survived CAI melting and slow cooling. Instead, we also interpret these assemblages as the products of low-temperature re-equilibration.

Our model for opaque assemblages predicts that the refractory siderophiles that are soluble in Ni-Fe metal at low temperature (such as Pt and Ir) should occur dissolved in Ni-Fe metal in opaque assemblages. This is generally the case, but occasionally Pt, Ir-rich nuggets are observed in sulphide-rich opaque assemblages. It has been shown^{11,12} that these nuggets formed during low-temperature sulphidization of Ni-Fe-Pt-Ir metal.

The occurrences of small (0.5–3 μm) refractory siderophile

metal nuggets in silicate or oxide phases in CAIs with no associated Ni-Fe metal, magnetite or sulphide^{7,10,13,37} are not specifically addressed by our model. These nuggets may be either primary condensate grains that did not come in contact with Ni-Fe metal droplets, remnants of digested opaque assemblages, or metals precipitated directly from the CAI silicate liquid.

Finally, it should be noted that isotopic studies of opaque assemblages and isolated refractory siderophile metal nuggets have found no evidence of large isotopic anomalies in any of the elements studied (Mg, Fe, Ru, Mo, W)³⁸. These results suggest that opaque assemblages, including Ru,Os-rich nuggets, were most probably formed within the solar nebula rather than outside the Solar System.

Conclusion

We have reproduced the basic assemblage of Ru,Os-rich nuggets, Ni-Fe metal and V-rich magnetite, characteristic of opaque assemblages, by oxidation of a homogeneous alloy at low temperature in the laboratory. We infer that opaque assemblages originated in CAI silicate liquid as homogeneous, probably molten, metal alloy droplets which included the elements Ru, Os, Re, Pt, Ir, W, Mo, P and V in the reduced state. The proto-CAI material most plausibly represents refractory early nebular condensates or evaporative residues. We propose that the complex mineralogy that now characterizes opaque assemblages formed from the alloy droplets by exsolution during progressive oxidation and sulphidization. This process probably occurred during the slow cooling of CAIs at a temperature of ~ 870 K. The alternative scenario, that opaque assemblages formed in the solar nebula before CAI melting and were subsequently captured in the CAIs^{9,11–13,15}, is not tenable because the low-temperature mineral assemblages could not have survived the CAI melting event for any plausible cooling rate. In our view, opaque assemblages do not contain much information on the high-temperature thermal histories of CAIs or record an independent history before CAI formation; they are not 'foreigners' to CAIs in the sense implied by the name "Fremdlinge". Instead, they are 'domestic' and provide equally important constraints on the lower-temperature thermal histories of CAIs and on the changing oxygen fugacity in cooling nebular and/or planetary environments.

We acknowledge the careful analytical and petrographic studies of "Fremdlinge" by J. T. Armstrong and A. El Goresy which set the framework for this study. We thank G. R. Rossman for making his laboratory available to us and H. Palme, K. Keil and the U.S. Museum of Natural History for loaning us the meteorite specimens used in this study, and H. Palme for a careful review. Funding was provided by grants from NASA and NSF, and by Caltech's Caroline Dobbins Fellowship to J.D.B. J.D.B., G.J.W. and I.D.H. are at the Lunatic Asylum of the Charles Arms Laboratory (No. 4537/591).

Received 19 October; accepted 24 December 1987.

- Grossman, L. A. *Rev. Earth planet. Sci.* **8**, 559–608 (1980).
- Wasserburg, G. J. & Papanastassiou, D. A. in *Essays in Nuclear Astrophysics* (eds Barnes, C. A., Clayton, D. D. & Schramm, D. N.) 77–140 (Cambridge University Press, New York, 1982).
- Grossman, L. *Geochim. cosmochim. Acta* **37**, 1119–1140 (1973).
- Wänke, H., Baddenhausen, H., Palme, H. & Spettel, B. *Earth planet. Sci. Lett.* **23**, 1–7 (1974).
- Fegley, B. & Kornacki, A. S. *Earth planet. Sci. Lett.* **68**, 181–197 (1984).
- Palme, H. & Wlotzka, F. *Earth planet. Sci. Lett.* **33**, 45–60 (1976).
- Wark, D. A. & Lovering, J. F. *Lunar Sci.* **7**, 912–914 (1976).
- El Goresy, A., Nagel, K., Dominik, B. & Ramdohr, P. *Meteoritics* **12**, 215–216 (1977).
- El Goresy, A., Nagel, K. & Ramdohr, P. *Proc. Lunar planet. Sci. Conf.* **9**, 1279–1303 (1978).
- Wark, D. A. & Lovering, J. F. *Lunar Sci.* **9**, 1214–1217 (1978).
- Armstrong, J. T., El Goresy, A. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **49**, 1001–1022 (1985).
- Armstrong, J. T., Hutcheon, I. D. & Wasserburg, G. J. *Geochim. cosmochim. Acta* (in the press).
- Bischoff, A. & Palme, H. *Geochim. cosmochim. Acta* **51**, 2733–2748 (1987).
- Fuchs, L. H. & Blander, M. *Proc. Lunar planet. Sci. Conf.* **11**, 929–944 (1980).
- Palme, H., Wlotzka, K., Nagel, K. & El Goresy, A. *Earth planet. Sci. Lett.* **61**, 1–12 (1982).
- Crank, J. *The Mathematics of Diffusion* (Oxford University Press, New York, 1975).
- Brandstätter, F. & Kurat, G. *Meteoritics* **18**, 272 (1983).
- Kracher, A. et al. *Proc. Lunar planet. Sci. Conf.* **16**, D123–D135 (1985).
- Swartzendruber, L. J. & Sundman, B. *Bull. Alloy Phase Diagrams* **4**, 155–160 (1983).
- Arrhenius, G. & Raub, C. J. *J. less-common Metals* **62**, 417–430 (1978).
- Blum, J. D., Armstrong, J. T., Hutcheon, I. D. & Wasserburg, G. J. *Lunar planet. Sci.* **18**, 97–98 (1987).
- Kaiser, T. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **47**, 43–58 (1983).
- MacPherson, G. J., Paque, J. M., Stolper, E. M. & Grossman, L. *J. Geol.* **92**, 289–305 (1984).
- Stolper, E. M. & Paque, J. M. *Geochim. cosmochim. Acta* **50**, 1785–1806 (1986).
- Wark, D. A. & Lovering, J. F. *Geochim. cosmochim. Acta* **46**, 2581–2594 (1982).
- Beckett, J. R. & Grossman, L. *Lunar planet. Sci.* **13**, 31–32 (1982).
- Barber, D. J., Martin, P. M. & Hutcheon, I. D. *Geochim. cosmochim. Acta* **48**, 769–783 (1984).
- Hirano, K., Cohen, M. & Averbach, B. L. *Acta Metall.* **9**, 440–445 (1961).
- Haggerty, S. E. & McMahon, B. M. *Proc. Lunar planet. Sci. Conf.* **10**, 851–870 (1979).
- McMahon, B. M. & Haggerty, S. E. *Proc. Lunar planet. Sci. Conf.* **11**, 1003–1025 (1980).
- Kelly, W. R. & Larimer, J. W. *Geochim. cosmochim. Acta* **41**, 93–111 (1977).
- Grossman, L., Olsen, E. & Lattimer, J. M. *Science* **206**, 449–451 (1979).
- Chipman, J. & Dastur, M. N. *J. Metals*, **N. Y.** **191**, 111–115 (1951).
- Palme, H. & Schmitt, W. *Lunar planet. Sci. Conf.* **17**, 623–624 (1986).
- Stolper, E. M., Paque, J. M. & Rossman, G. R. *Lunar planet. Sci.* **13**, 772–773 (1982).
- Beckett, J. R. & Grossman, L. *Lunar planet. Sci.* **17**, 36–37 (1986).
- Wark, D. A. *Earth planet. Sci. Lett.* **77**, 129–148 (1986).
- Hutcheon, I. D., Armstrong, J. T. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **51**, 3175–3192 (1987).
- Nash, P. *Bull. Alloy Phase Diagrams* **7**, 130–133 (1986).
- Kubaschewski, O. *Iron-Binary Phase Diagrams* (Springer, New York, 1983).
- Grossman, L. *Geochim. cosmochim. Acta* **36**, 597–619 (1972).
- Chase, M. W. Jr et al. *J. phys. chem. Ref. Data* **14**, Suppl. 1 (1985).

Temporal pattern of alcohol dehydrogenase gene transcription reproduced by *Drosophila* stage-specific embryonic extracts

Ulrike Heberlein & Robert Tjian

Howard Hughes Medical Institute, Department of Biochemistry, University of California, Berkeley, California 94720, USA

Regulation of transcription during development has been analysed in vitro using extracts derived from Drosophila embryos. Transcription of mutant templates reveals cis-control regions of the alcohol dehydrogenase promoter responsible for temporal regulation. The activity of a trans-acting protein correlates with the profile of alcohol dehydrogenase expression during embryogenesis.

THE molecular signals that trigger initiation of RNA synthesis in the *Drosophila* embryo are thought to be crucial in dictating the proper developmental pathway. Although molecular and genetic analyses of various *Drosophila* genes have identified *cis*-regulatory elements and *trans*-acting genes that participate in developmentally regulated expression *in vivo*¹⁻³, the molecular mechanisms governing temporally programmed patterns of transcription are not well understood. Insight into transcriptional regulation during *Drosophila* embryogenesis would be obtained from an *in vitro* system that recapitulates the transcriptional specificities intrinsic to embryos at various stages of development. Here we report the differential transcription of the *Drosophila melanogaster* alcohol dehydrogenase (Adh) gene in cell-free reactions using stage-specific embryonic extracts.

Adh is expressed at high levels in a tissue-specific⁴ and temporally regulated^{4,5} manner throughout the life cycle of the fly. The complex temporal regulation is achieved by the differential use of two tandemly arrayed promoters, distal and proximal, which are separated by ~700 base pairs (bp) (ref. 6), see Fig. 1. The transcript derived from the proximal promoter is first

observed during late embryogenesis. Transcription increases during the first and second larval instars, and falls abruptly in the middle of the third larval instar. The distal promoter is expressed under a different temporal programme, with low levels of transcript accumulating transiently in mid-stage embryos and later in third instar larvae. No Adh transcripts are found in pupae. Upon eclosion of adult flies, the level of distal transcript increases dramatically and remains high throughout most of the adult life. Small amounts of the proximal transcript are also observed in adult flies^{5,7}. The *cis*-acting DNA sequences which are sufficient to mediate temporal and tissue-specific control of Adh expression have been mapped by introducing altered Adh genes into flies by P-element-mediated transformation⁸. A region located within 660 bp upstream of the distal-transcription start site contains the *cis*-elements required for correct expression and regulation of this promoter. Sequences necessary for temporal and tissue-specific expression from the proximal promoter are found within 400 bp of the transcription start site, but other elements located ~2,000 bp upstream are apparently required for wild-type transcription levels.

We have previously used cell-free transcription reactions

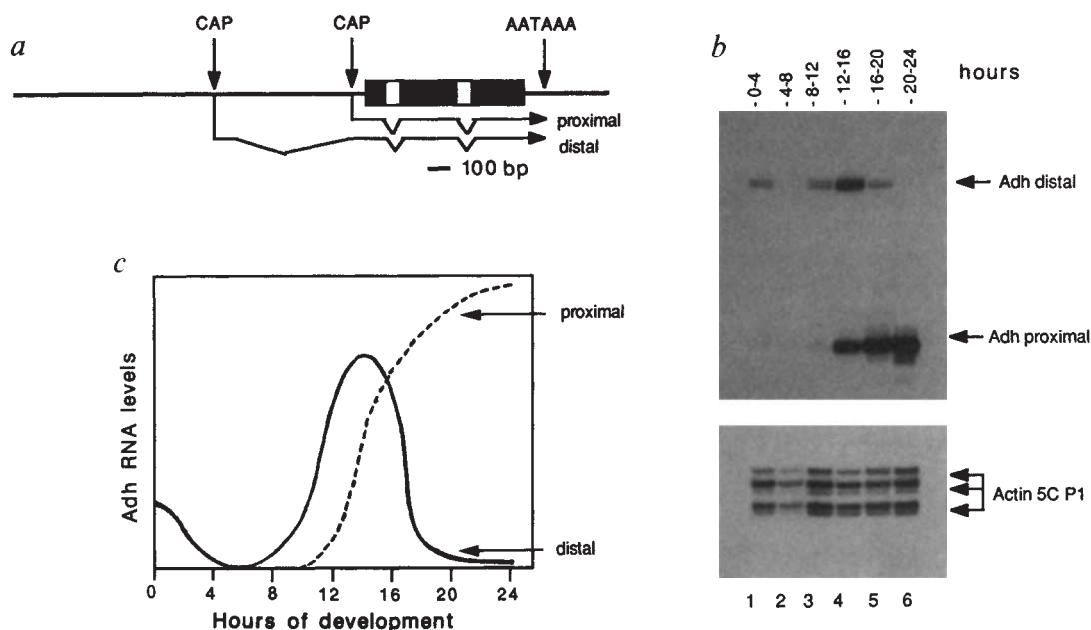


Fig. 1 Temporal pattern of *D. melanogaster* Adh RNA accumulation during embryogenesis. *a*, Structure of the distal and proximal Adh transcripts. The Adh coding regions are indicated as black boxes. *b*, Primer extension analysis of Adh RNA isolated from embryos at 4 h intervals from 0-24 h of age. 20 µg and 5 µg total embryo RNA (ref. 33) were analysed with the proximal Adh and actin 5C P1 primers respectively (for methods, see legend to Fig. 2). *c*, Schematic representation of data shown in *b*.

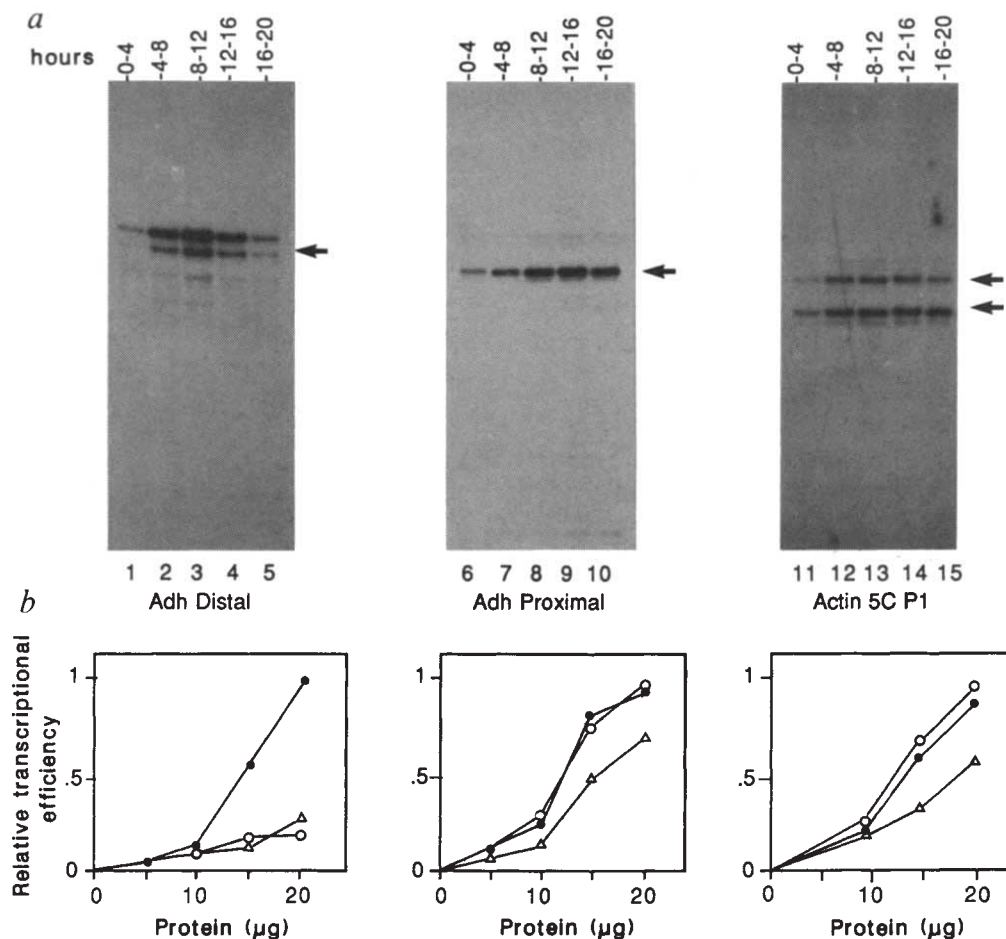


Fig. 2 *In vitro* transcription with stage-specific embryo extracts. **a**, Rates of transcriptional initiation of the Adh and actin 5C P1 promoters in extracts from embryos collected at various developmental stages were analysed by primer extension^{34,9}. Transcription of the distal and proximal Adh promoters was carried out simultaneously with a template, p13E-3 (see Methods), that contains both promoters in their normal arrangement. RNA samples were later divided and analysed with the distal (lanes 1-5) and the proximal (lanes 6-10) primers. The template used for actin 5C P1 transcription was pDmAz (ref. 35). All transcription reactions were carried out with 12 µg H.4 and 250 ng DNA template in 50 µl reaction volume. **b**, Relative transcription efficiencies of the Adh distal, proximal and actin 5C P1 promoters with different amounts of H.4 from early (2-7 h, Δ), mid (9-14 h, ●) and late (15-20 h, ○) embryos. *In vitro* transcription was assayed by primer extension and quantitated by densitometric scanning of autoradiograms.

Methods. Canton S wild-type flies were grown under standard conditions at 25 °C in population cages. Embryos were collected at 4 or 5 h intervals and either processed immediately or stored at 25 °C until reaching the appropriate age. Eggs were collected on molasses-agar trays covered with yeast paste, harvested and dechorionated by immersion in 2.25% sodium hypochlorite (1 in 2 dilution of commercial bleach) for 90 s. Dechorionated embryos were rinsed extensively with 0.7% NaCl, 0.04% Triton X-100, and finally with H₂O. Embryos were blotted dry and weighted. Between 10 and 20 g of embryos were used for the stage-specific nuclear extracts. All subsequent steps were carried out at 4 °C. Embryos were suspended in 2-4 ml g⁻¹ of embryo of buffer I (15 mM HEPES, pH 7.6, 10 mM KCl, 5 mM MgCl₂, 0.1 mM EDTA, 0.35 M sucrose, 1 mM dithiothreitol (DTT), 10 mM Na₂S₂O₅ and 1 mM PMSF), and homogenized with seven strokes of a motorized Teflon pestle (Wheaton) at 700 r.p.m. The homogenate was filtered through two layers of Miracloth (Calbiochem) pre-wetted with buffer I, and then centrifuged at 8,000 r.p.m. for 15 min in a Sorvall SS-34 rotor. The nuclear pellet was resuspended in 1 ml g⁻¹ of embryo of buffer II (15 mM HEPES, pH 7.6, 115 mM KCl, 5 mM MgCl₂, 0.1 mM EDTA, 1 mM DTT, 10 mM Na₂S₂O₅ and 1 mM PMSF) using a glass rod. Nuclei were lysed by adding 10% volume 4 M (NH₄)₂SO₄, pH 7.8, followed by slow mixing on a rotator (Scientific Instruments) for 30 min. The viscous extract was clarified by centrifugation at 35,000 r.p.m. for 1 h in a Beckman 60 Ti rotor. The supernatant was precipitated by the stepwise addition of 0.3 g ml⁻¹ of (NH₄)₂SO₄, and after stirring for 15 min, the precipitate was collected by centrifugation at 13,000 r.p.m. for 20 min in an SS-34 rotor. The pellet was carefully resuspended in 0.1 ml per g of embryo of buffer III (25 mM HEPES, pH 7.6, 40 mM KCl, 5 mM MgCl₂, 0.1 mM EDTA, 1 mM DTT and 10% glycerol), and dialysed against buffer III until the conductivity was equivalent to 100 mM KCl (~2.5-3 h). Precipitates were removed by centrifugation in an Eppendorf microfuge for 30 s. The protein concentration was typically between 30 and 40 mg ml⁻¹. Extracts were frozen in liquid N₂ and stored at -80 °C. The nuclear extracts were thawed and fractionated by heparin agarose chromatography. Typically, 30 mg protein were loaded on to a 1 ml column previously equilibrated with HEMG (25 mM HEPES, pH 7.6, 0.1 mM EDTA, 12.5 mM MgCl₂, 10% glycerol and 1 mM DTT) containing 0.1 M KCl (0.1 M KCl/HEMG). The column was washed with 0.1 M KCl/HEMG until the protein concentration in the flow-through was lower than 0.05 mg ml⁻¹. The bound protein was then eluted with 0.4 M KCl/HEMG. The fractions containing the eluate were pooled and dialysed against 0.04 M KCl/HEMG until the conductivity was equivalent to 100 mM KCl (protein concentration was 1-2 mg ml⁻¹). Aliquots were frozen in liquid N₂ and stored at -80 °C. This fraction, designated H.4, was used for both *in vitro* transcription and DNAase I footprinting experiments. *In vitro* transcription reactions and primer extension assays were carried out as described previously^{34,9}. Recombinant p13E-3 contains the 4.7 kilobase (kb) *Eco*RI fragment of gAcI (ref. 36) cloned into pUC-13. The primers used to analyse the distal and proximal Adh transcripts were complementary to sequences located between nucleotide positions +56 to +83 and +39 to +66, relative to the respective start sites. The primer for the actin 5C P1 transcript was complementary to nucleotides +43 to +67 relative to the transcription start site. The procedure to prepare nuclear extracts is a modification of a protocol obtained from W. Soeller and T. Kornberg, and was devised in collaboration with M. Biggin.

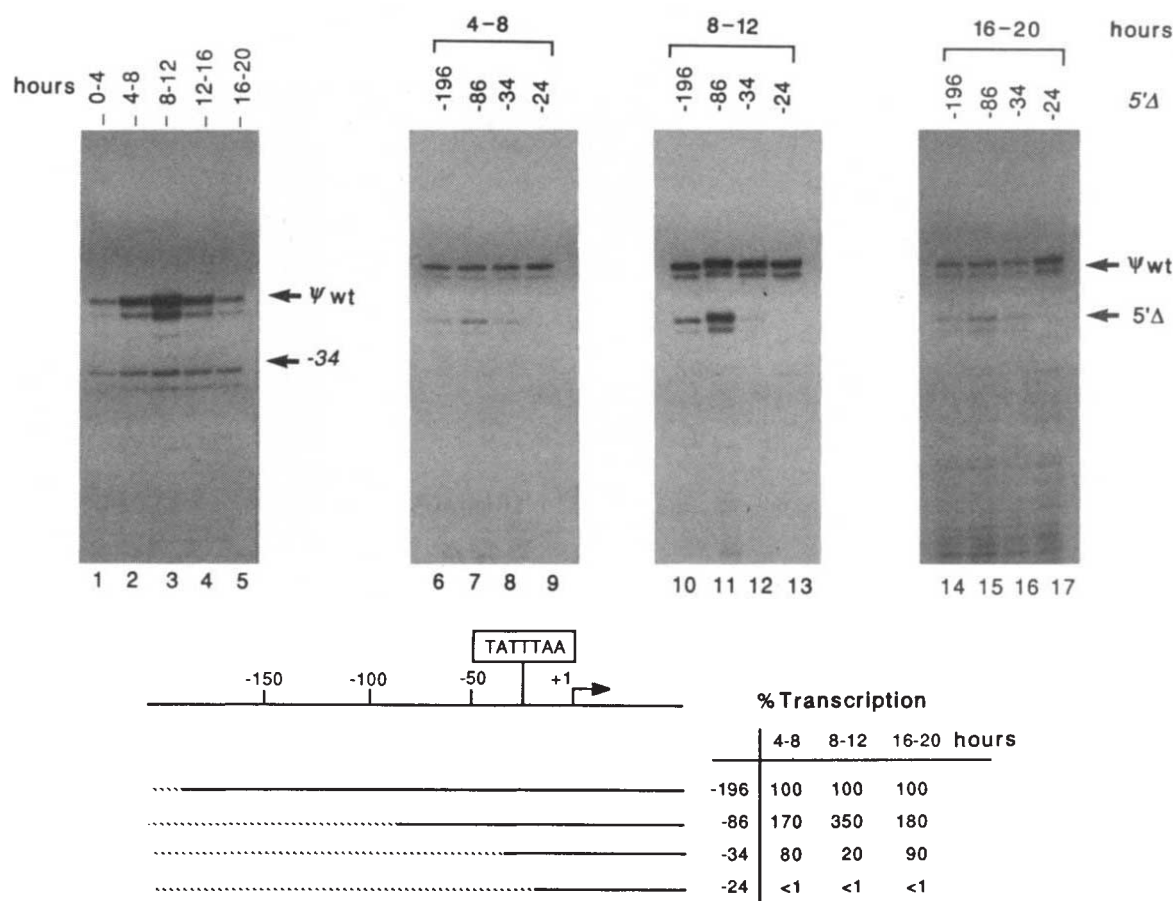


Fig. 3 Transcription of distal promoter mutants. *In vitro* transcription and primer extension analysis were carried out as described^{34,9}. The pseudo wild-type (ψ -wt) template included in the transcription reactions as an internal control contains an 8-bp insertion at +53 relative to the distal start site. The 5' deletion mutants have been described before⁹. Their 5' boundaries, relative to the start site of transcription, are as indicated and the 3' boundary is at about nucleotide +100. Reactions contained 15 μ g H.4 and 250 ng of each of the templates in 50 μ l final volume. The different panels cannot be compared directly with each other because they correspond to different autoradiographic exposures.

derived from *Drosophila* tissue culture cells to identify the DNA sequences and protein factors required for accurate and efficient initiation of Adh RNA synthesis *in vitro*⁹. Multiple upstream promoter elements were found to be required for optimal distal and proximal Adh promoter function. Fractionation of the nuclear extract and DNAase I footprinting showed that several DNA-binding proteins were present which interact specifically with DNA sequences required for efficient transcription *in vitro* and contained within the *cis*-regulatory regions required for proper regulation *in vivo*. One of these factors, Adf-1, binds to both Adh promoters and is necessary for maximal transcription from the distal, but not the proximal, promoter.

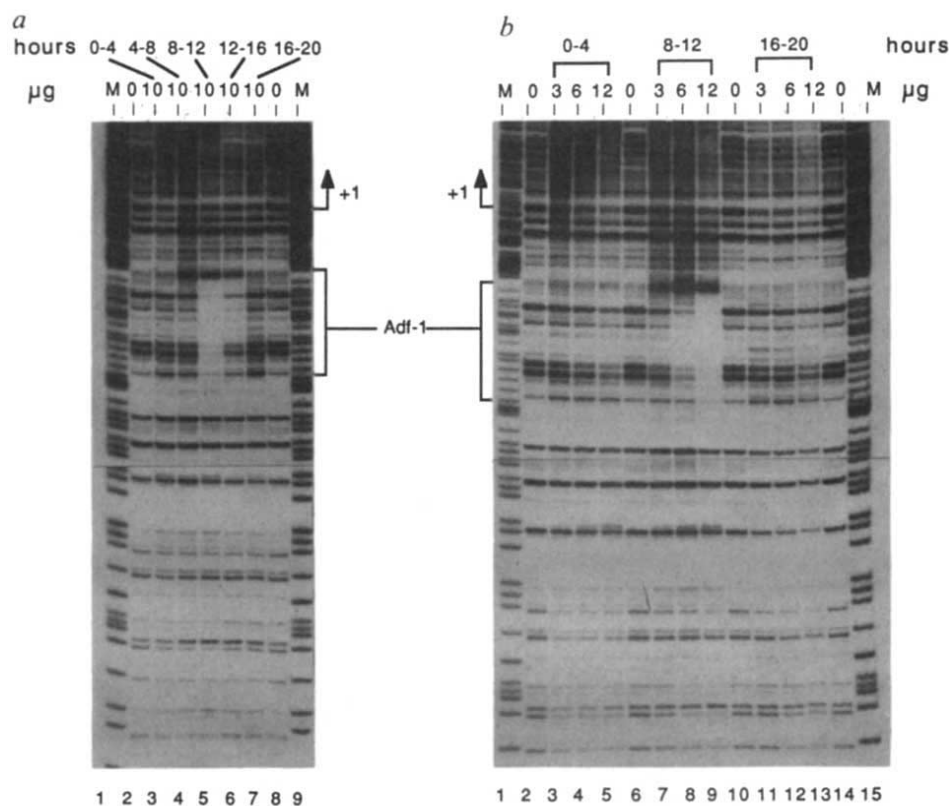
Kc cells are of embryonic origin¹⁰ but the cell type and developmental stage have not been identified. They provide a useful tissue culture system to study the basic mechanisms underlying initiation of transcription, but are less satisfactory for studying its temporal and/or spatial regulation. The preparation of extracts from *Drosophila* embryos has recently been reported¹¹. We have modified this procedure to prepare transcriptionally active extracts from specific stages of embryo development (see legend to Fig. 2) and to analyse other developmentally regulated *Drosophila* genes *in vitro* (M. Biggin and R.T., manuscript in preparation). Here we report that nuclear extracts from stage-specific embryos transcribe the distal Adh promoter with an activity profile that parallels the temporal regulation observed *in vivo* during embryogenesis. Using a combination of experiments, including promoter mutagenesis and

DNA-binding studies, we have mapped the *cis*-elements and identified a *trans*-acting protein factor involved in this regulation *in vitro*. Interestingly, we find that the binding activity of Adf-1 correlates with the transcriptional activity of the distal promoter in stage-specific embryo extracts. We also show that purified Adf-1 is able to complement extracts deficient in endogenous Adh transcriptional activity.

In vitro transcription

In an attempt to develop homologous systems to study the temporal regulation of transcription in *Drosophila*, a series of nuclear extracts from staged embryos were prepared, each extract representing a different period of embryonic development. One series of five extracts was derived from embryos collected at 4-h intervals between 0–20 h of development. A second series of three extracts was obtained from early, mid-stage and late embryos, corresponding to ages 2–7, 9–14 and 15–20 h respectively. We prepared total RNA from a portion of each collection and determined the amounts of endogenous Adh RNA as a standard of comparison for the transcriptional activities *in vitro*. An example, in which the levels of Adh RNA were measured at 4-h intervals, is shown in Fig. 1b, and schematically represented in Fig. 1c. These results, obtained by primer extension of the Adh RNA, are very similar to those obtained using Northern blots⁵ or RNAase protection assays⁷ in that the early embryos contained low amounts of the distal and the proximal transcript, mid-stage embryos had high

Fig. 4 DNAase I footprint analysis of transcription factor Adf-1 activity in stage-specific embryo extracts. *a*, Binding reactions were carried out as described⁹ with 0 (lanes 2 and 8) and 10 μ g H.4 derived from the staged embryo extracts as indicated (lanes 3–7). The distal promoter probe was end-labelled at position –240 (ref. 9). Calf thymus DNA, 1 μ g in 50 μ l reaction volume, was used as non-specific competitor. *b*, Footprinting conditions were the same as in *a*. Increasing amounts of H.4 from embryos of ages 0–4 h (lanes 3–5), 8–12 h (lanes 7–9), and 16–20 h (lanes 11–13) were used. Digestion ladders in the absence of protein are indicated with 0. Maxam and Gilbert sequencing markers³⁷ (M) are displayed alongside the DNAase-digestion ladders.



levels of both, and only the proximal RNA was detected in late embryos.

Plasmids containing the Adh promoters or the actin 5C P1 (ref. 12) promoter were used as templates to direct *in vitro* transcription reactions with the developmental stage-specific extracts. These extracts were first fractionated by heparin-agarose chromatography (see legend to Fig. 2) to enrich for active transcription factors. All components necessary for accurate and efficient initiation of transcription elute between 0.1 M and 0.4 M KCl (H.4 fractions) and were used in the experiments presented here.

The H.4 fractions derived from embryos at all the different developmental stages support accurate initiation of transcription at the distal and proximal Adh promoters. However, the efficiency of transcription directed by both promoters varies greatly between extracts derived from embryos at different stages of development (Fig. 2). Transcription of the wild-type distal Adh promoter (Fig. 2a, lanes 1–5) is low in extracts from 0–4-h-old embryos. Activity increases with the age of the embryos until reaching a peak in the extract from 8–12-h-old embryos, corresponding to an 8–10-fold increase in the rate of initiation. Extracts from older embryos transcribe the distal promoter with significantly lower efficiency: a 6–8-fold reduction is observed in extracts from 16–20 h compared to 8–12-h-old embryos. Thus, the pattern of transcription from the distal promoter *in vitro* resembles the temporal profile observed for the transcript *in vivo* during embryogenesis (compare Fig. 1b and Fig. 2a, lanes 1–5). Although the maximal rate of distal transcription *in vitro* was observed with extracts from 8–12-h-old embryos, the peak of distal transcript abundance *in vivo* was usually centred around 12–14 h of development. We do not understand the nature of this difference, but one explanation is that our *in vitro* transcription reactions measure the rate of initiation of RNA synthesis, whereas the RNA observed *in vivo* corresponds to accumulation of steady-state levels of the transcript.

Transcription of the proximal Adh promoter with the stage-specific embryo extracts follows a temporal pattern quite different from that of the distal promoter (Fig. 2a, lanes 6–10).

Activity is low in extracts from 0–4- and 4–8-h-old embryos, and increases 3–5-fold in extracts from 8–12-h-old embryos, but the transcription from the proximal promoter remains high in extracts from late embryos. Likewise, *in vitro* transcription of the actin 5C P1 promoter with the stage-specific extracts is relatively constant, with the exception of the extract from 0–4-h-old embryos, which is about one third as active (Fig. 2a, lanes 11–15). Measurement of the transcription efficiency over a range of extract concentrations (see Fig. 2b) confirms the differential activity of the staged embryo extracts. These findings indicate that nuclear extracts derived from *Drosophila* embryos at different stages of development transcribe the distal Adh promoter with a temporal pattern strikingly similar to that observed *in vivo*.

The transcriptional efficiency of all promoters tested is lower in the extracts from early embryos, and in particular from 0–4-h-old embryos. Addition of these extracts to the more active ones caused no reduction in transcription, and addition of a fraction containing RNA polymerase II from Kc cells did not stimulate transcription (data not shown). Lower activity is thus not due to the presence of specific inhibitors or to limiting amounts of RNA polymerase II. Very low rates of RNA synthesis by RNA polymerase II have been observed *in vivo* during the early stages of embryonic development^{13–16}. These observations could explain our results with extracts from early embryos. Whereas the low transcriptional competence of extracts from early embryos is general to all promoters tested, extracts from late embryos show reduced activity with only the distal Adh promoter. A series of mixing experiments showed that this decreased activity does not result from the presence of inhibitors that selectively reduce transcription from the distal promoter (data not shown). Therefore it is likely that factors specifically required for efficient initiation at this promoter are absent or depleted in extracts from late embryos.

Mapping of *cis*-regulatory DNA sequences

As a first step towards defining the DNA sequences responsible for distal promoter transcription and temporal regulation *in*

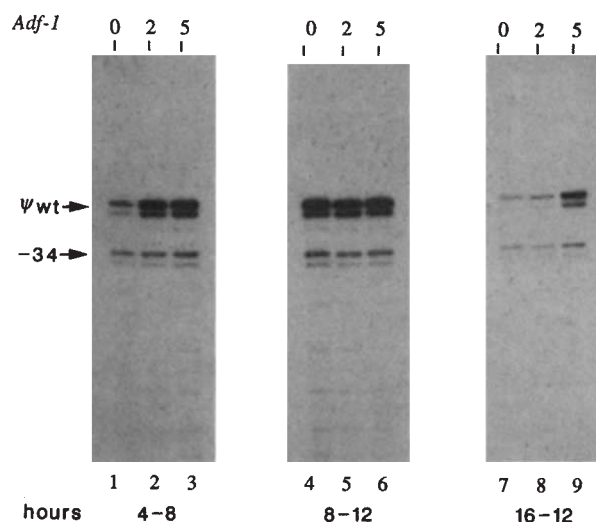


Fig. 5 Effect of purified Adf-1 on distal promoter transcription by stage-specific embryo extracts. The transcription reactions contained 250 ng pseudo wild-type template, 250 ng of the template deleted to -34 , 10 μ g H.4 from embryos of age 4–8 h (lanes 1–3), 8–12 h (lanes 4–6) and 16–20 h (lanes 7–9) and increasing amounts (in μ l) of Adf-1 (~ 5 ng μ l $^{-1}$), as indicated in the figure. Transcription and primer extension analysis were carried out as described^{24,9}.

vitro, we have used the stage-specific embryo extracts to transcribe templates lacking different portions of upstream promoter sequences. We found that a pseudo wild-type template containing a 200-bp sequence of the distal promoter is transcribed with the same temporal specificity and relative efficiency as a template containing an upstream sequence of 1,500 bp. In contrast, a distal promoter retaining only 34 nucleotides of 5' sequence is transcribed to a lesser extent which remains constant, regardless of the stage-specific extract used (Fig. 3, lanes 1–5). This result indicates that the sequences mediating temporal control *in vitro* are probably located between nucleotide positions -200 and -34 .

We further dissected this *cis*-regulatory region by analysing a series of 5'-deletion mutants, with end-points located between -200 and -24 relative to the transcription start site. We tested their ability to support specific initiation (see Fig. 3, lanes 6–17). Each transcription reaction also contained the pseudo wild-type template as an internal control. Deletion of sequences located between nucleotide positions -196 and -86 results in roughly a twofold increase in transcription with extracts from 4–8- and 16–20-h-old embryos, and a 3–4-fold increase with extracts from 8–12-h-old embryos. Transcription of the template with a deletion to position -34 is reduced by 50% with extracts from 4–8- and 16–20-h-old embryos, and to about 6% with extracts from 8–12-h-old embryos. Deletion of sequences located between -34 and -24 , including the TATA box, essentially abolishes specific initiation in all extracts tested. This deletion mutant analysis suggests that at least 3 *cis*-elements are influencing distal Adh transcription. The element located between nucleotides -196 and -86 appears to affect transcription negatively, whereas the regions between -86 and -34 , and -34 and -24 are necessary for efficient initiation of RNA synthesis at this promoter. Our analysis also indicates that the region responsible for the temporal regulation of the distal promoter *in vitro* is contained within the sequences between nucleotide positions -34 and -86 relative to the start site of transcription.

Binding activity of transcription factor Adf-1

We have previously reported the presence in *Drosophila* Kc cells of a sequence-specific DNA-binding factor Adf-1, which binds to and activates transcription from the distal Adh promoter in an *in vitro* transcription system⁹. Because the Adf-1-binding site, located between nucleotides -46 and -86 , lies within a region involved in temporal regulation of transcription, we tested stage-specific embryo extracts for the presence of Adf-1 using DNAase I footprinting¹⁷. We investigated the correlation of this factor with the differential transcription observed (Fig. 4): H.4 fractions derived from 8–12-h-old embryos contain large amounts of Adf-1, whereas protein from 0–4- and 16–20-h-old

embryos shows significantly decreased binding (Fig. 4a). Extracts from 4–8- and 12–16-h-old embryos give weak Adf-1 footprints, corresponding to intermediate amounts of bound factor. To quantitate these differences in Adf-1 activity, we compared the degree of protection from DNAase I obtained after incubating the distal promoter probe with increasing amounts of embryo H.4. From the data presented in Fig. 4b, we conclude that early and late embryo extracts contain about a quarter of the Adf-1 activity of extracts from 8–12-h-old embryos. Using poly d(I-C), instead of calf thymus DNA, as a non-specific competitor in the footprinting reactions, we detected several other DNA-binding factors which specifically interact with the distal promoter sequences located between -160 and -92 , and between -42 and $+90$ (unpublished observation). Interestingly, the activities of these other factors in the stage-specific embryonic extracts do not follow the same temporal profiles as Adf-1. Most of these DNA-binding activities are found predominantly in late embryos (unpublished observations), whereas Adf-1 binding is maximal in extracts from mid-stage embryos. Thus we find that Adf-1, a positive transcription factor for the distal promoter, binds to sequences required for stage-specific transcription of this promoter. Moreover, changes in binding activity follow a temporal pattern that parallels transcription of the distal promoter in stage-specific extracts.

Effect of Adf-1 on transcription

If the decreased efficiency of transcription of the distal promoter in extracts from early and late embryos is a result of limiting amounts of Adf-1, transcription should be stimulated in these extracts by the addition of exogenous factor. Adf-1 has been purified to apparent homogeneity from *Drosophila* Kc cells and embryos (B. England and R.T., in preparation) using sequence-specific DNA-affinity chromatography¹⁸ by a method described for the transcription factor Sp1 (ref. 19). Figure 5 shows an experiment in which purified Adf-1 has been added to the transcription reactions carried out by the stage-specific embryo extracts. A template containing an Adf-1-binding site (ψ -wt), together with an internal control template deleted to nucleotide position -34 , and therefore lacking the Adf-1 binding site, was transcribed in the presence of increasing amounts of purified Adf-1. Addition of Adf-1 to extracts from 4–8- and 16–20-h-old embryos, which contain low endogenous amounts of the factor, stimulated transcription sixfold (Fig. 5 lanes 1–3 and 7–9). Adding the same amount of purified factor to the extract from 8–12-h-old embryos, which already contains high levels of Adf-1, did not significantly stimulate transcription (Fig. 5, lanes 4–6). These results indicate that Adf-1 is indeed a limiting factor in the rate of transcription with early and late embryo extracts.

Discussion

Several lines of evidence indicate that the previously characterized cellular transcription factor Adf-1 is important in the temporal regulation of the distal Adh promoter *in vitro*. First, the *cis*-control sequences that mediate differential activation of the distal Adh promoter in staged-embryo extracts contain a binding site for Adf-1. Second, the sequence-specific DNA-binding activity of Adf-1 in embryos at different stages of development parallels the transcriptional activity of the Adh promoter *in vivo* and *in vitro*. Finally, purified Adf-1 protein is able to supplement extracts deficient in endogenous distal transcriptional activity, such as those derived from early and late embryos. These findings show that cell-free extracts prepared from staged *Drosophila* embryos provide a useful system for identifying and isolating *trans*-regulatory factors involved in establishing specific transcription patterns during embryogenesis. This approach has also been successful in studying transcription of the *Drosophila* ultrabithorax gene *in vitro* (M. Biggin and R.T., in preparation).

Although Adf-1 seems to be important in regulating transcription from the distal Adh promoter *in vitro*, it is probably not the only factor involved. Using different non-specific competitor DNAs in our footprinting reactions, we detected several other DNA-binding factors which specifically interact with the distal promoter elements that modulate *in vitro* transcription (unpublished results). Thus, the distal promoter of the *Drosophila* Adh gene appears to be composed of a complex arrangement of *cis*-regulatory modules recognized by multiple, distinct, *trans*-acting proteins, a situation established for numerous mammalian promoters²⁰⁻²². Further studies are required to establish the influence of these different factors on Adh transcription.

Our experiments indicate that extracts from early and late embryos contain lower Adf-1-binding activity than extracts from mid-stage embryos, but the reason for this is unclear. Amounts of Adf-1 were comparable in extracts from whole embryos and in nuclear extracts (data not shown), suggesting that differences in Adf-1 activity are not a result of changes in its cellular localization or of selective loss of nuclei containing active Adf-1. We cannot exclude the possibility that Adf-1 is inactivated during the preparation of late or early extracts, but the presence of other DNA-binding factors in the same extracts suggests that any inactivation must be selective for Adf-1. We do not know what mechanisms regulate Adf-1 activity during embryogenesis: the amount of Adf-1 protein present could vary during development or, by analogy with the inducible systems in

mammalian²³⁻²⁶ and *Drosophila*²⁷ cells, Adf-1 could be inactive in early and late embryos and activated by covalent modification during embryogenesis. Raising antibodies against the factor and characterizing the gene(s) encoding it should help elucidate the mechanisms underlying the changes in Adf-1 activity. To determine the importance of Adf-1 in Adh transcription during the development of the fly, we have recently generated germ-line transformants which contain Adh genes with the Adf-1 site on the distal and/or proximal promoters deleted. Our preliminary analysis of these transformants indicates that distal transcription in embryos and adult flies is reduced by mutations in the Adf-1 binding site, suggesting that Adf-1 is involved in regulating distal promoter function *in vivo*. Deletion of the Adf-1 binding site on the proximal promoter had no effects on the production of this transcript.

Although promoter specificity and regulation can be the result of coordinate action of different *trans*-regulators²⁰⁻²², it is also possible that a specific factor could have distinct functions in regulating a variety of promoters. Thus, a factor such as Adf-1 could confer a different range of transcriptional activity to different promoters. Therefore we tested other *Drosophila* promoters for potential Adf-1 responsive elements: antennapedia P1 (M. Biggin, B. England and K. Perkins, personal communication) and the dopa decarboxylase promoters were found to bind Adf-1 with an affinity comparable to the Adh promoters (unpublished observations). Transcripts originating at these promoters, however, follow different temporal patterns of accumulation during embryonic development²⁸⁻³². Thus, the presence of Adf-1 binding sites alone is not sufficient to confer a particular temporal pattern of expression. This is consistent with the observation that both the distal and proximal promoter interact with Adf-1, but display distinct temporal patterns of transcription. Analysis of these promoters *in vivo* and *in vitro*, together with the biochemical characterization of the *trans*-acting protein factors that recognize them, will provide further insight into the complex mechanisms governing temporal control of transcription during *Drosophila* development.

We thank Katie Rendahl and Gina Hess for help with *Drosophila* culture, Bruce England for gifts of Adf-1, Jim Merryweather for synthesis of oligonucleotides, and Karen Ronan for help with the figures. We are grateful to Mark Biggin, Karen Perkins, Bruce England, Craig Hauser, Frank Laski, Drzilav Mismar and Michael Jantzen for critical reading of this manuscript. This work was supported in part by a grant from NSF to R.T.

Received 2 November; accepted 30 December 1987.

1. Scott, M. P. & O'Farrell, P. H. *A. Rev. Cell Biol.* **2**, 49-80 (1986).
2. Gehring, W. J. *Science* **236**, 1245-1252 (1987).
3. Maniatis, T., Goodbourn, S. & Fischer, J. A. *Science* **236**, 1237-1245 (1987).
4. Ursprung, H., Sofer, W. H. & Burroughs, N. *Wilhelm Roux Arch. dev. Biol.* **164**, 201-208 (1970).
5. Savakis, C., Ashburner, M. & Willis, J. H. *Dev. Biol.* **114**, 194-207 (1986).
6. Benyajati, C., Spoerel, N., Haymerle, H. & Ashburner, M. *Cell* **33**, 125-133 (1983).
7. Fisher, J. A. & Maniatis, T. *EMBO J.* **5**, 1275-1289 (1986).
8. Posakony, J. W., Fischer, J. A. & Maniatis, T. *Cold Spring Harbor Symp. quant. Biol.* **50**, 515-520 (1986).
9. Heberlein, U., England, B. & Tjian, R. *Cell* **41**, 965-977 (1985).
10. Echalié, G. & Ohanessian, A. C. *R. Acad. Sci. (Paris), Ser. B* **268**, 1771-1773 (1969).
11. Soeller, W. C., Poole, S. J. & Kornberg, T. *Genes Dev.* (in the press).
12. Bond, B. J. & Davidson, N. *Molec. cell Biol.* **6**, 2080-2099 (1986).
13. Zolotar, M. *Dev. Biol.* **49**, 425-437 (1976).
14. Anderson, K. V. & Lengyel, J. A. *Dev. Biol.* **70**, 217-231 (1979).
15. McKnight, S. L. & Miller, O. L. Jr *Cell* **8**, 305-319 (1976).
16. Edgar, B. A. & Schubinger, G. *Cell* **44**, 871-877 (1986).
17. Galas, D. & Schmitz, A. *Nucleic Acids Res.* **5**, 3157-3170 (1978).
18. Kadonaga, J. T. & Tjian, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5889-5893 (1986).
19. Briggs, M. R., Kadonaga, J. T., Bell, S. P. & Tjian, R. *Science* **234**, 47-52 (1986).
20. Dynan, W. S. & Tjian, R. *Nature* **316**, 774-778 (1985).
21. McKnight, S. & Tjian, R. *Cell* **46**, 795-805 (1986).
22. Serfling, E., Jasin, M. & Schaffner, W. *Trends Genet.* **1**, 224-230 (1985).
23. Sen, R. & Baltimore, D. *Cell* **47**, 921-928 (1986).
24. Pyrwes, R. & Roeder, R. G. *Cell* **47**, 777-784 (1986).
25. Hayes, T. E., Kitchen, A. M. & Cochran, B. H. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1272-1276 (1987).
26. Kingston, R. E., Schuetz, T. J. & Larin, Z. *Molec. cell Biol.* **7**, 1530-1534 (1987).
27. Zimarino, V. & Wu, C. *Nature* **327**, 727-730 (1987).
28. Laughon, A., Boulet, A. M., Bermingham, J. R. Jr, Laymon, R. A. & Scott, M. P. *Molec. cell Biol.* **6**, 4676-4689 (1986).
29. Schneuwly, S., Kuriowa, A., Baumgartner, P. & Gehring, W. J. *EMBO J.* **5**, 733-739 (1986).
30. Marsch, J. L. & Wright, T. R. F. *Dev. Biol.* **80**, 379-387 (1980).
31. Hirsch, J. & Davidson, N. *Mol. cell Biol.* **1**, 475-485 (1981).
32. Kraminsky, G. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4175-4179 (1980).
33. O'Hare, K., Lewis, R. & Rubin, G. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6917-6921 (1983).
34. McKnight, S. L. & Kingsbury, R. *Science* **217**, 316-324 (1982).
35. Fyrberg, E. A., Kindle, K. L., Davidson, N. & Sodja, A. *Cell* **19**, 365-378 (1980).
36. Goldberg, D. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5794-5798 (1980).
37. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).

Gamma-ray line emission from SN1987A

S. M. Matz*, G. H. Share*, M. D. Leising*,
E. L. Chupp†, W. T. Vestrand†, W. R. Purcell‡,
M. S. Strickman* & C. Reppin§

* E. O. Hulburt Center for Space Research, Naval Research Laboratory, Washington, DC 20375, USA

† Department of Physics, University of New Hampshire, Durham, New Hampshire 03824, USA

‡ Department of Physics and Astronomy, Northwestern University, Evanston, Illinois 60201, USA

§ Max Planck Institute for Physics and Astrophysics, D-8046 Garching, FRG

The gamma-ray spectrometer (GRS) on NASA's Solar Maximum Mission satellite (SMM) has observed a significant ($>5\sigma$) net line flux at ~ 847 keV in the background-subtracted spectrum of SN1987A—accumulated between 1 August and 31 October 1987. This is the energy of the strong gamma-ray line from decay of ^{56}Co , which was predicted¹⁻³ to be seen from supernovae. The inferred average line flux during this period is $\sim (1.0 \pm 0.25) \times 10^{-3}$ photons $\text{cm}^{-2} \text{s}^{-1}$ at an energy of 843 ± 5 keV. This feature cannot be explained by any statistical or systematic fluctuations observed in the seven previous years of GRS data. There is also evidence for the 1,238-keV line from ^{56}Co decay, with an average flux of $\sim (6 \pm 2) \times 10^{-4}$ photons $\text{cm}^{-2} \text{s}^{-1}$. The quoted photon fluxes for both lines are preliminary. This observation confirms that ^{56}Co is present in the supernova ejecta (as implied by the light curve^{4,5}) and that nucleosynthesis took place during the explosion. The first appearance of the gamma-ray lines was roughly coincident with the first detections of X-rays by Ginga and MIR^{6,7}.

The GRS⁸ consists of seven 3-inch diameter by 3-inch thick NaI detectors, actively shielded on the sides and back by CsI. With rare exceptions, the satellite and instrument are pointed at the Sun. SN1987A is therefore (with a variation of a few degrees) at 90° to the forward viewing direction, and must be observed through the side CsI shield. For each orbit, a source spectrum is obtained by accumulating the data taken when the supernova was unocculted by the Earth. A corresponding background spectrum is produced by summing the occulted data. Subtracting these two spectra essentially eliminates the background features due to radioactivity in the spacecraft and instrument, but the difference will still show a net flux due to unequal exposure to atmospheric gamma radiation in the source and background accumulations. The atmospheric residual is fairly smooth in the energy ranges of interest.

We have specifically examined these background-subtracted spectra for evidence of gamma-ray line emission from the decay of ^{56}Co . To obtain the line intensities, the background-subtracted counting rate spectra are fitted in the regions near the expected lines with a combination of a gaussian peak and a power-law continuum. The line energies are fixed at the values for a source at rest, and the widths at the instrumental resolution. The amount of Doppler broadening expected from the expansion of the ejecta would not be observable in the GRS. Figure 1 shows the intensities of the 847-keV line resulting from fits to (approximately) 10-day sums of the background-subtracted spectra, over seven years of SMM operation. Because only an *a posteriori* data selection is used to make the source/background subtraction, we can produce spectra of the supernova region for years before the explosion. There are periodic gaps in the history because, due to orbital inclination and precession, there are intervals when the Earth does not occult the source. Between the time of the supernova explosion and about the end of July the fitted intensities are consistent with the intensities fitted

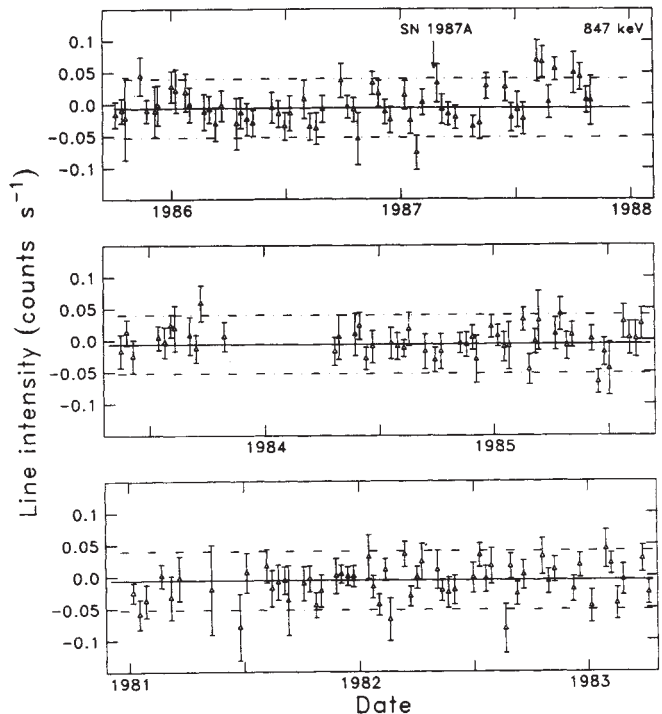


Fig. 1 The time history of the 847-keV line intensity from fits to the background-subtracted counting rate spectra. The data have been divided into three equal intervals, with the earliest times at the bottom and the latest times at the top. The gap around day 1500 is due to missing data in the period before satellite repair in April 1984. The mean intensity of the data points before 1 August is indicated by a solid line. For reference, intensities of ± 0.046 counts s^{-1} from the mean are shown by dashed lines. The time of the supernova explosion is indicated by the arrow.

before the explosion. Since the beginning of August an increase in flux is evident at 847 keV. This period is represented by the last eight points in Fig. 1.

The statistical significance of this increase in the 847-keV line intensity can be derived from the data in Fig. 1. The data before the last eight points are consistent, at the 20% level, with a random distribution about a mean intensity of $(-6.1 \pm 1.9) \times 10^{-3}$ counts s^{-1} , based on χ^2 per degree of freedom (d.o.f.) = 1.095 for 153 degrees of freedom. Including the final eight points raises the $\chi^2/\text{d.o.f.}$ to 1.296, and reduces the chance that the overall distribution is random to 1.2%. The mean of the last eight points is $(4.0 \pm 0.9) \times 10^{-2}$ counts s^{-1} . The probability (from χ^2) that the last eight points are a statistical fluctuation from the earlier mean is $\sim 6 \times 10^{-6}$. This probability must be increased by the number of points since the supernova explosion, giving a final value of $\sim 7 \times 10^{-5}$.

The top spectrum in Fig. 2 shows the background-subtracted counting rate spectrum accumulated over the period of increased flux (1 August to 31 October). The solid line indicates the peak positions and instrumental widths for the two ^{56}Co decay lines at 847 and 1238 keV. Also shown is a spectrum accumulated under similar background conditions in 1985. Note the residual atmospheric continuum in both spectra. The feature near 847 keV in the top spectrum has an intensity of $(3.7 \pm 0.7) \times 10^{-2}$ counts s^{-1} with the peak energy and width fixed at the nominal values. The best-fit energy of the peak is 843 ± 5 keV, in good agreement with the expected energy; the line width is 59 ± 14 keV, consistent with the instrumental width of 57 keV.

The inferred average photon flux in the 847 keV line from 1 August to 31 October is $\sim (1.0 \pm 0.25) \times 10^{-3}$ photons $\text{cm}^{-2} \text{s}^{-1}$,

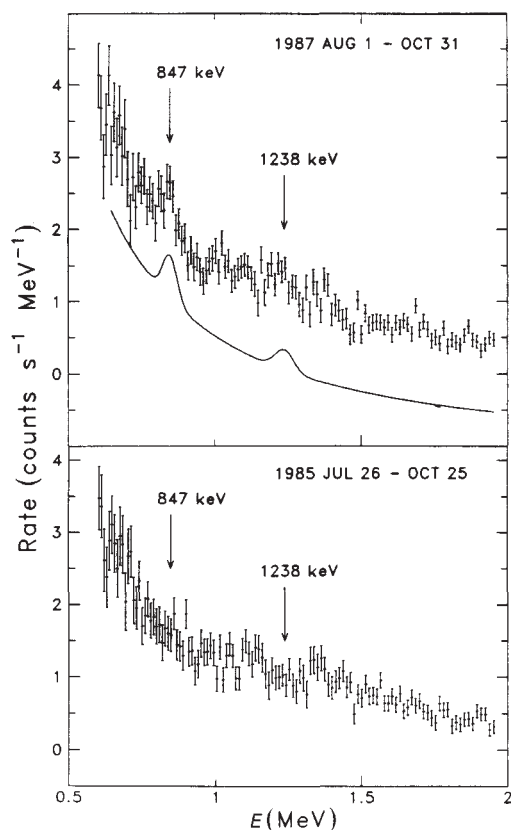


Fig. 2 The background-subtracted spectrum for SN1987A, accumulated from 1 August to 31 October 1987. The line profiles for the two ^{56}Co lines plus a power-law continuum are plotted as a solid line. Also shown is an equivalent spectrum accumulated in 1985. In both spectra residual features can be seen near 1.17 and 1.33 MeV due to imperfect subtraction of lines from the internal calibration source. Note also the residual atmospheric continuum in both spectra.

assuming that the line comes from the supernova. The quoted error in the flux has been increased over the statistical error to include an estimate of possible systematic contributions to the measured line intensity. This estimate is based on background-subtracted spectra accumulated during times of similar background and occultation conditions in previous years. The spectrum for 1985 in Fig. 2 is one such accumulation. None of these spectra show any significant net flux at 847 keV. Since no systematic 847-keV line is observed in these pre-supernova spectra, we have used the observed range of fitted fluxes as an estimate of the possible systematic contribution to the line intensity in 1987. This is added to the statistical error from the fit to give an estimated total error on the inferred photon flux. Note again that this error does not reflect the statistical significance of the line detection itself, which is greater than 5σ .

The spectrum and the significance of the increase in line intensity indicate that the observed feature is unlikely to be a statistical fluctuation. The possibility remains that it is a real line produced locally, either in the Earth's atmosphere or by activation of spacecraft material. There is no observed atmospheric line at this energy, and the appearance of the 847-keV line is not correlated with the residual atmospheric continuum in the spectra. There are possible lines near this energy from the decay of radioactive isotopes produced in the spacecraft, notably ^{27}Mg and ^{56}Mn . We have therefore made tests to determine if the 847-keV line is associated with activation due to increases in cosmic ray background (measured by vertical cutoff rigidity) or passage through the radiation belts in the South Atlantic Anomaly (SAA). Data taken at times of high rigidity (low background) show the same intensity (within statistics) of

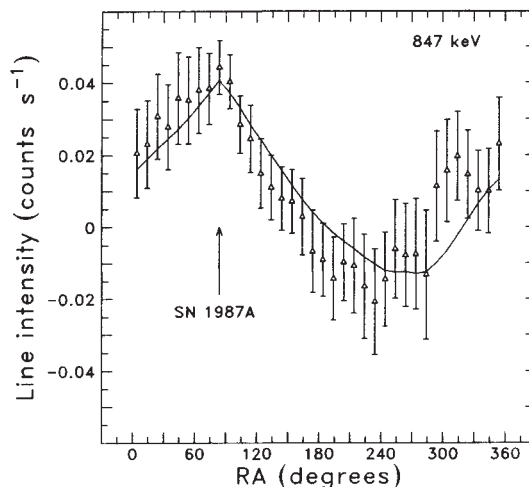


Fig. 3 The fitted 847-keV line intensities for spectra accumulated for test positions differing from the position of the supernova in right ascension. Also shown are the intensities expected assuming that the line emission originates from SN1987A. The right ascension of the supernova is indicated by an arrow.

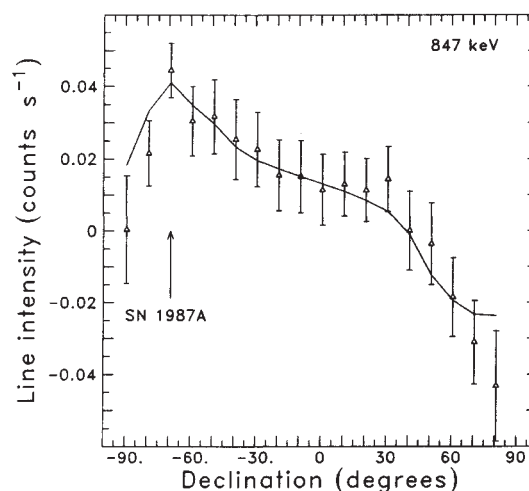


Fig. 4 The fitted 847-keV line intensities for spectra accumulated for test positions differing from the position of the supernova in declination. Also shown are the intensities expected assuming that the line emission originates from SN1987A. The declination of the supernova is indicated by an arrow.

the 847-keV line as data taken at low rigidity, as do data collected at least two ^{27}Mg half-lives from times of low rigidity. There is no correlation of line intensity with difference in average rigidity between source and background either before or after the supernova. All data used were taken more than 3 h from the last significant SAA passage, eliminating contributions from radioactives with substantially shorter half-lives. The subtraction done orbit-by-orbit eliminates contributions to the difference from long half-life isotopes. Increasing the lower limit on time from last SAA does not affect the 847-keV line intensity. The line intensity does not correlate with the difference in average time from last SAA between the source and background spectra. The values of rigidity and time from SAA observed from 1 August to 31 October are consistent with those seen at earlier times.

As another test for systematics, we have used the data from 1 August to 21 October to produce spectra for celestial positions differing from that of the supernova by rebinning the data into

source and background according to the occultations of the new position. If the observed line flux is actually due to emission from SN1987A, the measured flux at a different assumed position will depend only on the relative exposure to the supernova in the source and background spectra for that position. Figures 3 and 4 show the fitted fluxes for positions differing in right ascension and declination from that of the supernova. Also shown are the expected fluxes based on exposure. In both cases the data are in good agreement with the supernova exposure. Note however that the data points are not independent, so the error bars are misleading. The variations of intensities of apparent background features in the spectrum do not agree well with the model.

We have looked for a second strong line of ^{56}Co decay at 1,238 keV. Based on the branching ratio, the 1,238 keV line should be present with at least 67.9% of the 847-keV flux, or 6.8×10^{-4} photons $\text{cm}^{-2} \text{s}^{-1}$. If there is significant attenuating material above the ^{56}Co , this flux should be higher, perhaps exceeding that of the 847 keV line (see discussion below). Imperfect subtraction of the lines from the internal calibration sources may leave positive or negative features at 1,173 and 1,333 keV. These are visible in the two spectra in Fig. 2. Fitting the 1,238-keV line simultaneously with the residual background lines gives a flux of $(6 \pm 2) \times 10^{-4}$ photons $\text{cm}^{-2} \text{s}^{-1}$ at $1,218 \pm 16$ keV. This intensity is consistent with the lowest expected flux. We have applied the systematic check using test positions described above to the 1,238-keV line. The 1,238-keV line intensity variations agree with those expected for emission from the supernova. However, both the statistics and the systematics make this result less certain than the detection of the 847-keV line.

The fluxes in these lines are equivalent to what would have been observed during this period if there were $\sim 2.3 \times 10^{-4} M$ of totally exposed ^{56}Co present at 55 kpc on 1 August. This is only $\sim 1.3\%$ of the total mass of ^{56}Co thought to be present in the supernova ejecta at that time, based on the light curve^{4,5}. This suggests that the observed gamma rays may have been produced by a small fraction of the total ^{56}Co under very little material, or by all the ^{56}Co under a thick attenuating envelope. Since 847-keV photons are more likely to scatter than 1,238-keV photons, a thick envelope will enhance the 1,238-keV flux with respect to that at 847 keV. The observed line ratio is $F_{1,238}/F_{847} = 0.60 \pm 0.25$, consistent with the laboratory branching ratio of 68%. This places a limit on the thickness of the material above the ^{56}Co . If the observed gamma-ray lines were produced by the total mass of ^{56}Co under a thick attenuating envelope, the overlying material would have to have an average effective optical depth of $\tau_{847} \approx 4$ to reduce the 847-keV flux to the measured intensity. A Monte Carlo calculation⁹ indicates that an envelope this thick would produce a line ratio of 1.1, which is unlikely, but not excluded by the data. Another problem with this model is that some of the scattered gamma rays will emerge as X-rays, and the calculated hard X-ray flux is about ten times that actually observed by MIR^{6,9}.

Alternatively, the observed gamma rays may have been produced by a small fraction of the total mass of ^{56}Co under very little material. Such a situation might arise in the supernova if the envelope is non-uniform, or if a small amount of ^{56}Co has moved or been mixed out beyond the bulk of the ejecta. In this case, the measured line ratios will be close to the branching ratio. In addition, the mass of material overlying the observed ^{56}Co can be chosen to make the gamma-ray fluxes consistent with the observed hard X-rays. This would, however, produce a spectrum somewhat harder than is observed⁹. It is interesting to note that one explanation proposed for the luminosity of the mysterious "companion" of the supernova requires that a similar mass of ^{56}Co be expelled as a blob or jet^{10,11}.

The GRS has observed a line at an energy consistent with 847 keV in the background-subtracted spectra for SN1987A for over 80 days. No previously observed statistical or systematic fluctuations can explain this feature. We conclude that this line

is probably due to the decay of ^{56}Co in the supernova ejecta. Although we cannot absolutely rule out that this is the first appearance of a new and unexpected background effect, we believe that this is unlikely to be the case. There is also evidence for the presence of the 1,238-keV decay line. The line fluxes do not appear to have risen since the initial detection. The best fit to the data indicates that the flux is slowly declining, but the data are consistent with a wide range of models, including a constant or slowly rising intensity. Based on the best fit to the time history since August, we would predict that the balloon experiments launched last autumn would have seen a flux of roughly $(3-7) \times 10^{-4}$ photons $\text{cm}^{-2} \text{s}^{-1}$ in the 847 keV line. GRS observations of SN1987A are continuing.

We are indebted to Dave Forrest who had major responsibility for the design, development, and calibration of the instrument. Jim Kurfess provided critical evaluation of the analysis. Techniques used in the analysis were derived from algorithms developed by Bob Kinzer and Neil Johnson. We thank Dan Messina for his direction of the production analysis and careful screening of the data, and Augie Iadicicco and James Johnson for assisting in data processing. Jim Ryan and Mike Cherry made important contributions to the development and calibration of the instrument. We also thank our colleagues working on the other SMM experiments, and the personnel at Goddard Spaceflight Center responsible for mission planning, operation, and data retrieval, for their assistance in making these observations. This work was supported by NASA at NRL and UNH, and by BMFT at Max-Planck Institut. S.M.M. and M.D.L. are National Research Council Resident Research Associates at the Naval Research Laboratory.

Received 4 January; accepted 19 January 1988.

1. Clayton, D. D., Colgate, S. A. & Fishman, G. J. *Astrophys. J.* **155**, 75-82 (1969).
2. Bodansky, D., Clayton, D. D. & Fowler, W. A. *Astrophys. J. Suppl. Ser.* **16**, 299-371 (1968).
3. Colgate, S. E. & McKee, C. *Astrophys. J.* **157**, 623-643 (1969).
4. Catchpole, R. M. *et al. Mon. Not. R. Astr. Soc.* (submitted).
5. Woosley, S. E. *Astrophys. J.* (submitted).
6. Sunyaev, R. *et al. Nature* **330**, 227-230 (1987).
7. Dotani, T. *et al. Nature* **330**, 230-231 (1987).
8. Forrest, D. J. *et al. Solar Phys.* **65**, 15-23 (1980).
9. Leising, M. D. *Nature* (submitted).
10. Goldman, I. *Astr. Astrophys.* **186**, L3-L4 (1987).
11. Piran, T. & Nakamura, T. *Nature* **330**, 28 (1987).

Observation of an ionospheric disturbance caused by a gamma-ray burst

G. J. Fishman* & U. S. Inan†

* Space Science Laboratory, NASA/Marshall Space Flight Center, Huntsville, Alabama 35812, USA

† Space, Telecommunications and Radioscience Laboratory, Stanford University, Stanford, California 94305, USA

We report a first observation of an ionospheric disturbance from a gamma-ray burst. The burst, GB830801, occurred at 22:14:18 UT on 1 August 1983 and was one of the strongest ever observed. The total fluence was 2×10^{-3} erg cm^{-2} , most of which occurred in the first 4 s of the burst. Simultaneously, a change was observed in the amplitude of a very-low-frequency (VLF) radio signal from a transmitter in Rugby, England, monitored at Palmer Station, Antarctica, indicative of an ionospheric disturbance. Weaker disturbances were also recorded at the same receiving site on signals from VLF stations in Annapolis, Maryland and Lualualei, Hawaii. The times of the burst and the disturbances are coincident within the 10-s resolution of the VLF recording system. No similar disturbances were observed within 60 h around the time of the burst. In the future, a network of VLF burst monitors may provide

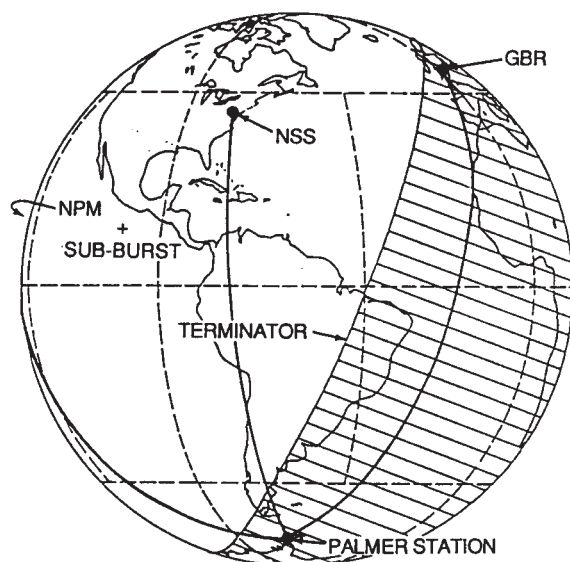


Fig. 1 The great-circle paths of the three VLF signals are shown along with the location of the sunset terminator at the time of the gamma-ray burst. The sub-burst position is indicated by a cross.

measurements of the total ionizing energy fluence from a burst, as well as some limited directional information.

VLF radio propagation is a sensitive probe of the lowest portions of the ionosphere, from 40 km to 90 km¹. In this region cosmic rays provide the major source of ionization. VLF radio signals travel long distances with little attenuation and their propagation is best modelled by wave-guide mode theory^{2,3}. High-power continuous VLF radio transmissions are used by several countries for global navigation and communication. An enhancement of ionization in the lower ionosphere can be detected as an amplitude change or a phase shift of the received VLF signal. The magnitude and the sign of the amplitude change are dependent on the transmitter-receiver path, the VLF frequency and the altitude profile of the ionization change^{4,5}.

In several fields of research, VLF propagation has been used to monitor transient ionospheric enhancements (disturbances). For many years, solar flares have been observed through sudden ionospheric disturbances⁶ (SIDs). Monitoring distant VLF radio signals has been shown to be one of the most sensitive indicators of hard X-ray and microwave emission from flares⁶⁻¹⁰. Bain and Hammon¹⁰, for example, noted the success in detecting solar flares by VLF-phase-anomaly monitoring for solar X-ray flares with peak intensity $>6 \times 10^{-4} \text{ erg cm}^{-2} \text{ s}^{-1}$ (3–20 keV). A threshold for detection of $2 \times 10^{-3} \text{ erg cm}^{-2} \text{ s}^{-1}$ was reported by Kreplin *et al.*⁷. VLF propagation studies are also used as probes of sporadic electron precipitation caused by disturbances in the Earth's magnetosphere^{4,5,11-15}. One type of transient precipitation is believed to be induced by whistlers which, in turn, are generated by strong lightning discharges¹⁶⁻¹⁸. The monitoring of VLF propagation has also been used to study the effects on the ionosphere of strong X-ray sources such as Sco X-1, Cyg X-3, Cen X-2 and Cen X-4, but the results have been

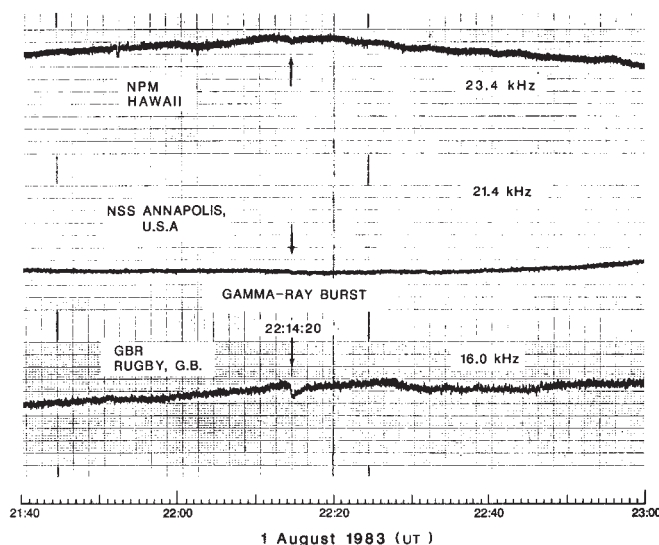


Fig. 2 Amplitude (arbitrary units) of the VLF radio signals received at Palmer, Antarctica from the three stations indicated. The ionospheric disturbances at the time of the gamma-ray burst are indicated by arrows.

unconfirmed¹⁹⁻²⁴. Hudson and TeKolste²⁵ have suggested monitoring VLF propagation to detect pulsed ionization enhancements from strong X-ray pulsars such as A0535+26 and the Crab Pulsar.

The effects of a gamma-ray burst on the ionosphere were first calculated by Brown²⁶. Others have also published calculations of the ionization profiles expected from gamma-ray bursts^{23,28-30}. These calculations show that the peak of the ionization due to a gamma-ray burst takes place in the region from 25 km to 35 km, but significant ionization occurs, primarily from Compton electrons, up to what is usually considered the nighttime reflection height for VLF propagation, $\sim 85 \text{ km}$ ²⁷.

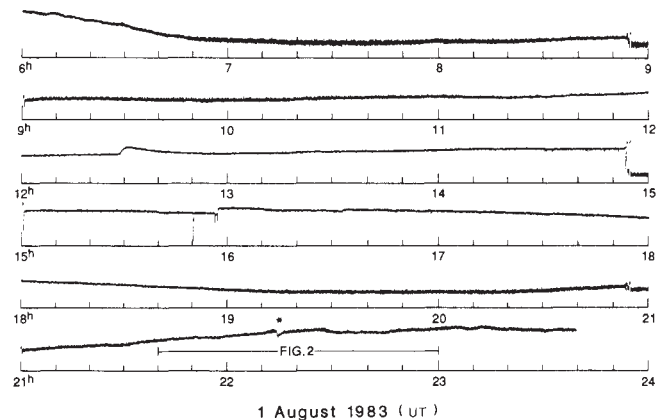
The observations reported here were obtained with apparatus normally used to observe the relationships between whistlers and their effects on VLF propagation¹⁶. The receiving station was located at the Palmer Station, Antarctica (65° S, 64° W). An analog strip chart recorder monitored various sources of data relevant to the magnetospheric studies, with a time resolution of about 10 s. Three channels consisted of the amplitude of the received signal from distant VLF transmitters (Fig. 1). Some properties of the transmitters and the propagation distances are given in Table 1.

Figure 2 shows a portion of the record from the three stations between 21:40 UT and 23:00 UT on 1 August 1983. A clear indication of a disturbance beginning at 22:14:10 $\pm$ 10 UT is seen in the radio station GBR signal. Weaker, barely detectable decreases in amplitude are seen simultaneously in the two other signals. Without the GBR signal, these other two signals alone would have been considered uneventful as similar weak fluctuations are seen in their records near the time of the burst. The disturbance in the GBR signal differs in its rise-and-fall time from any other disturbances seen within 60 h of the burst. The GBR disturbance has a rise time of $<20 \text{ s}$ and returns to the

Table 1 VLF transmission paths

Station	Location	Frequency (kHz)	Latitude	Longitude	Path to Palmer, Antarctica Arc length	Distance ($\times 10^6 \text{ m}$)
GBR	Rugby, UK	16.0	52°24' N	01°12' W	127°	14.1
NPM	Lualuaie, Hawaii	23.4	21°25' N	158°09' W	107°	11.9
NSS	Annapolis, USA	21.4	38°59' N	76°27' W	105°	11.6

Fig. 3 Extended record (18 h) of the 16-kHz signal from station GBR. A more typical solar-type disturbance is seen at 12:40 UT, having a longer rise and fall time than the disturbance due to the gamma-ray burst at 22:14 UT. Other sudden jumps in the record are due to adjustments in the transmitter or in the receiving equipment. The time of the gamma-ray burst is indicated by a star.



pre-disturbance baseline in ~ 200 s. A 16-h sample of the GBR signal is shown in Fig. 3.

The gamma-ray burst at 22:14:20 UT, 1 August 1983 (GB830801), was observed by the Signe-2MP9 experiment on the Prognoz-9 satellite^{31,32}, by the International Comet Explorer³³ (ICE) and by the Vela spacecraft (J. Laros, personal communication). The burst was one of the strongest recorded to date, with a total energy fluence of at least 2×10^{-3} erg cm⁻² (ref. 31). It was observed over the energy range from 5 keV to 7.5 MeV. In addition to its intensity, the burst was unusual because of its smooth time profile. The higher energy radiation rose to near maximum in 1 s, then dropped to $\sim 15\%$ of the peak in 5 s. The lower energy radiation lasted for >40 s (ref. 33).

The observed ionospheric disturbance is attributed to the gamma-ray burst GB830801 due to its temporal coincidence and the fact that solar X-ray flares of this intensity level are observed to produce disturbances of similar magnitude^{6,10}. An important distinction between solar flares and gamma-ray bursts is that the typical flare duration is longer and the mean photon energy is lower. Furthermore, recent quantitative models of ionization due to sporadic electron precipitation and its effect on VLF radio propagation would support the hypothesis that a gamma-ray burst with a total energy of 10^{-3} erg cm⁻² would produce the observed effect^{4,5}.

The gamma-ray burst location was derived from timing measurements from three widely separated spacecraft: ICE, Prognoz and Vela. The most likely location was RA=11 h 48 min, dec. = +13°, in the constellation Leo, with an error radius of $\sim 5^\circ$ (J. Laros, personal communication). At the time of the burst, the point on the Earth directly beneath the burst (sub-burst point) was 103° W longitude +13° S latitude, ~ 800 km southwest of Guatemala. Fortunately, all three of the transmission paths (Fig. 1, Table 1) were entirely within the hemisphere of the Earth that was irradiated by the burst. The path from Rugby to Palmer, Antarctica, was 65° to 85° from the sub-burst point, and the other two transmission paths were considerably closer. The fact that the disturbance measured by the GBR signal was greater than that of the other two signals may be due to this signal path occurring almost entirely during night-time. The VLF reflection height for a night-time ionosphere is typically at ~ 85 km (ref. 27). If the amplitude change is due to increased absorption, then a night-time path would allow the signal to traverse a region of greater integrated ionization. At the time of the burst, no solar activity was reported in the radio or microwave region³⁴.

There are several possible uses for the observation of gamma-ray bursts through ionospheric disturbances. It may be possible to localize better some burst directions through a large network of observations. Although the zenith angle dependence of ionization due to a burst is rather small²⁶, observations from a network of various path lengths could indicate the fraction of the paths that are in the irradiated hemisphere of the burst. For other observing conditions being equal, the magnitude of the disturbance should be proportional to the path length affected by the burst. These observations could also yield the total ionizing

fluence of a gamma-ray burst. This measurement is difficult to obtain from spaceborne observations due to the limited energy range of the detectors used for gamma-ray burst observations and because of detector saturation in the observation of the strongest bursts. For those bursts which can be measured by both space-borne and VLF techniques, they would provide a calibration of the magnitude of VLF propagation disturbances from known ionizing sources (M. Walt, personal communication). These data would be important for continuing studies of magnetospheric-ionospheric coupling processes.

Based on the known log N -log S distribution of strong gamma-ray bursts³⁵ (spectral index = -1.5), and assuming that the observed signal was about five times the minimum detectable signal, then only a few gamma-ray bursts per year can be studied through ionospheric disturbances. However, coincidence techniques and a network of receiving sites could considerably improve the observability of burst-induced disturbances.

As a final note of interest, this may be the first time that a transient extra-solar phenomenon has measurably affected a part of the Earth's environment.

We are grateful for helpful discussions with R. L. Showen, D. L. Carpenter, H. S. Hudson, J. Laros and K. Hurley. The Stanford work is supported by the NSF.

Received 13 October; accepted 6 November 1987.

- Sechrist, C. S. Jr *Radio Sci.* **9**, 137-149 (1974).
- Watt, A. D. *VLF Radio Engineering* (Pergamon, Oxford, 1967).
- Kikuchi, T. *J. atmos. terr. Phys.* **48**, 15-23 (1986).
- Tolstoy, A., Rosenberg, T. J., Inan, U. S. & Carpenter, D. L. *J. geophys. Res.* **91**, 13473-13482 (1986).
- Tolstoy, A. *thesis Univ. Maryland* (1983).
- Mitra, A. P. *Ionospheric Effects of Solar Flares* (Reidel, Dordrecht, 1974).
- Kreplin, R. W., Chubb, T. A. & Friedman, H. *J. geophys. Res.* **67**, 2231-2253 (1962).
- Kaufmann, P. & Paes de Barros, M. H. *Sol. Phys.* **9**, 478-486 (1969).
- Kaufmann, P. & Rizzo Piazza, L. *Sol. Phys.* **57**, 479-481 (1978).
- Bain, W. C. & Hammon, E. *J. atmos. terr. Phys.* **37**, 573-574 (1975).
- Potemra, T. A. & Rosenberg, T. J. *J. geophys. Res.* **78**, 1572-1580 (1973).
- Westerlund, S. & Reder, F. H. *J. atmos. terr. Phys.* **35**, 1453-1474 (1973).
- Helliwell, R. A., Katsufakis, J. P. & Trimpi, M. L. *J. geophys. Res.* **78**, 4679-4688 (1973).
- Muraoka, Y. *J. atmos. terr. Phys.* **41**, 1031-1042 (1979).
- Kikuchi, T. & Evans, D. S. *J. geophys. Res.* **88**, 871-880 (1983).
- Inan, U. S., Carpenter, D. L., Helliwell, R. A. & Katsufakis, J. P. *J. geophys. Res.* **90**, 7457-7469 (1985).
- Inan, U. S. & Carpenter, D. L. *J. geophys. Res.* **92**, 3293-3303 (1987).
- Goldberg, R. A., Barcus, J. R., Hale, L. C. & Curtis, S. A. *J. atmos. terr. Phys.* **48**, 293-299 (1986).
- Ananthakrishnan, S. & Ramanathan, K. *Nature* **223**, 488-489 (1969).
- Svennesson, J. & Westerlund, S. *J. atmos. terr. Phys.* **41**, 361-365 (1969).
- Svennesson, J., Reder, F. & Crouchley, J. *J. atmos. terr. Phys.* **34**, 49-72 (1972).
- Kaufmann, P., Piazza, L. R. & Ananthakrishnan, S. *Astrophys. Space Sci.* **22**, 67-70 (1973).
- Kasturirangan, K., Rao, U. R., Sharma, D. P., Rastogi, R. C. & Chakravarty, S. C. *Astrophys. Space Sci.* **42**, 57-62 (1976).
- Rizzo Piazza, L., Kaufmann, P. & Ramirez Pardo, P. *J. atmos. terr. Phys.* **45**, 121-125 (1983).
- Hudson, H. & TeKolste, R. *Bull. Am. astr. Soc.* **16**, 982 (1984).
- Brown, R. T. *Nature* **246**, 83-84 (1973).
- Deeks, D. G. *Proc. R. Soc. Ser. A*, **291**, 413-425 (1966).
- Baird, G. A. *J. atmos. terr. Phys.* **36**, 1565-1569 (1974).
- O'Mongain, E. & Baird, G. A. *Astrophys. Space Sci.* **42**, 63-67 (1976).
- Weekes, T. C. *J. atmos. terr. Phys.* **38**, 1021-1026 (1976).
- Kuznetsov, A. V. *et al. Sov. Astr. Lett.* **12**, 315-318 (1986).
- Boer, M. *et al. Adv. Space Res.* **6**, 97-102 (1986).
- Laros, J. in *Gamma Ray Bursts* (eds Liang, E. P. & Petrosian, V.) 84-86 (AIP Press, New York, 1986).
- Solar-Geophysical Data No. 474 Part II*, 6 (NOAA, Boulder, 1984).
- Mazets, E. *et al. Astrophys. space Sci.* **80**, 3-83 (1981).

Diameter of the Sun in AD 1715

Leslie V. Morrison*, F. Richard Stephenson†
& John Parkinson‡

* Royal Greenwich Observatory, Herstmonceux Castle, Hailsham, East Sussex BN27 1RP, UK

† Department of Physics, University of Durham, Durham DH1 3LE, UK

‡ Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK

Ribes *et al.*¹ have reopened the discussion about the variability of the Sun's diameter over the past few hundred years by asserting that observations made at the Paris Observatory by Picard and La Hire² during the late seventeenth and early part of the eighteenth century indicate that the diameter of the Sun was ~ 4 arc s greater than it is now. In two earlier papers^{3,4} the conclusion drawn from analyses of solar eclipses and transits of Mercury was that there has been little or no discernible secular change in the Sun's diameter since 1715. Ribes *et al.*¹ dismissed this result and concluded that their result based on La Hire's observations, and ours^{3,4} based on the 1715 eclipse, were not incompatible. Here we say that the two results are incompatible, and we give a new and careful discussion of the 1715 eclipse in support of our contention.

It is a fortunate accident of nature that the Moon and Sun have nearly equal diameters as seen from the surface of the Earth. This concurrence produces total eclipses of the Sun of relatively short duration (typically a few minutes), and this affords a method of measuring the Sun's diameter, with the Moon—whose mean diameter and profile are well known—acting as a template.

Due to the brilliance of the photosphere, the beginning and end of an eclipse are very well-defined events which can be observed easily without the aid of a telescope.

Where observers are situated at the extreme northern and southern limits of the path of totality, timings are not required to deduce the diameter of the Sun. We simply require to know where the observers were and that they saw the Sun diminish to a point of light before increasing again, or disappearing momentarily. Several such reports of the 1715 eclipse were collected by Halley⁵. Figure 1 shows the positions of the principal observers and the boundaries of totality in England for the 1715 eclipse. The oval shadow moved from south-west to north-east between the boundaries at a speed of $\sim 3,000$ km h⁻¹. Throughout our calculations we have used the modern value of 959.63 arc s for the semi-diameter of the Sun at unit distance.

The southern limit is fixed by two independent witnesses whose observations were reported by Halley⁵. Pages 256 and 257 of Halley's account are reproduced here photographically (Fig. 2). The crucial observations were made at Bocton, where there was a small light left on the lower part of the Sun that appeared like a star; and Cranbrook, where the Sun was extinguished for a moment, and instantly emerged again. Figure 3c, d shows the calculated limits of totality in relation to Cranbrook and Bocton. Muller⁶ consulted local historians in Cranbrook who informed him that William Tempest resided at Angley House on the outskirts of Cranbrook (see Fig. 3c). The position angle of contact as seen from both places was near 180° and from charts of the lunar profiles⁷ the depth of the relevant lunar depression is found to be -1.8 arc s. Hence, the position of the computed southern limit is displaced 1.8×2.6 km = 4.6 km, approximately north-west. The factor 2.6 is the change in the limit in kilometres produced by a change of 1 arc s in the Sun's semi-diameter. It is this computed limit which is plotted in Fig. 3a–d and labelled '959.63'. Figure 3a, b shows the southern limit in two areas of Sussex relevant to the discussion below.

The crucial observation on the northern limit (Fig. 4) was

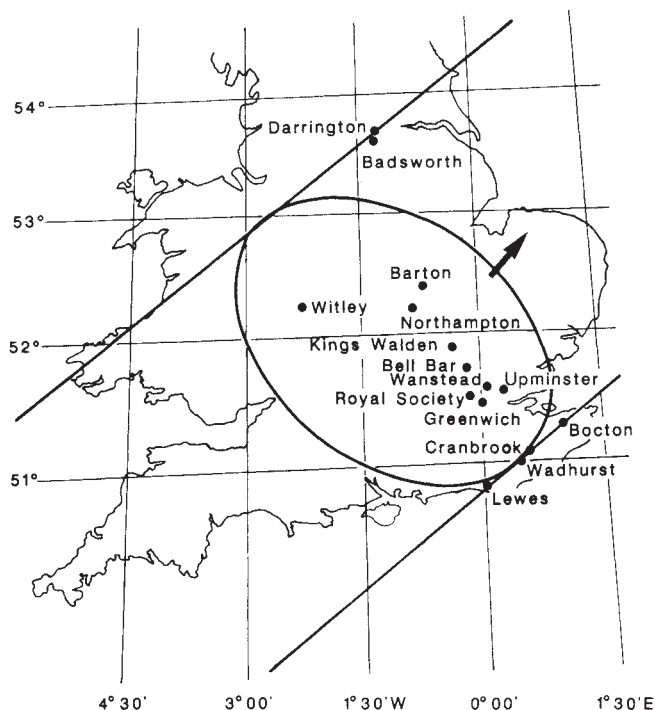


Fig. 1 Path of totality of 1715 eclipse, showing the position of the oval shadow at one instant. Observations were made from the places marked on the map.

made at Darrington, where Shelton observed the Sun to be reduced in size, such that it resembled the planet Mars, before increasing again. The limb profile correction near position angle 0° is -0.6 arc s. The calculated northern limit, therefore, is moved 0.6×2.6 km = 1.6 km, approximately south-east. This corrected limit is drawn in Fig. 4 and labelled '959.63'.

From Fig. 3c, d it can be seen that the southern limit was observed ~ 0.3 km south-east of the calculated limit, whilst from Fig. 4 it can be seen that the northern limit was observed 0.5 km to the south-east of the calculated limit. These results imply a shift of the central line of the eclipse path south-east by half the sum of these amounts (0.4 km) and a correction to the semi-diameter of the Sun of half the difference (-0.05 km ≈ -0.1 arc s). A shift of 0.4 km in the central line would arise from a change of +1 second of time to the ephemeris longitude which is quite plausible for the early part of the eighteenth century⁸.

Hence, from the 1715 total solar eclipse we find the correction to the modern value (959.63 arc s) of the semi-diameter to be ≈ -0.1 arc s from the northern and southern limits of totality.

An increase of ~ 4 arc s in the solar diameter as contended by Ribes *et al.*¹ would have produced a narrower path of totality, moving the northern and southern limits closer together by 4×2.6 km = 10.4 km; or 5.2 km at each limit. These limits are labelled '961.63' in Fig. 3 and 4, assuming an increase of exactly 2 arc s in the semi-diameter. The eclipse would then not have been total at Lewes and Wadhurst near the southern limit (see Fig. 3a, b), nor Badsworth near the northern limit (see Fig. 4), in contradiction to the affirmative reports⁴ (Fig. 2, p. 256, and 258 which begins "And that it was so (Total) at Badsworth about the same Distance from Darrington, we are told by a Letter of the Reverend and Learned Mr Daubuz, that he has a certain Account from that Place, that the luminous Ring round the Moon was seen there, which was no where visible but while the Eclipse was Total.").

This new result from the path width of the 1715 eclipse, and our earlier result³ from the timed duration at the places shown in Fig. 1, lead us to the conclusion that the semi-diameter of the Sun in 1715 was not significantly different from the present-day value (959.63 arc s at unit distance). Ribes *et al.*¹ must have

(256)

and R. S. S. assisted by that accurate Observer Mr Stephen Gray; by which we learn that the Eclipse began there at 8^h. 8'. 55" and ended at 10^h. 24'. 47"; and that the Total Darkness continued but about one Minute or rather less, the middle thereof being at 9^h. 13'. 52". From this Duration it will follow that Norton-court was but about 3 or 4 Miles within the Shade. And that it was really so is confirmed by the Relation of the Inhabitants of Bocton, about Midway between Norton-court and Canterbury, who assured Mr. Gray, as he was returning home that same Day, that the Eclipse was not Total there, but, as one of them exprest it, before the Sun had quite lost his Light on the East-side he recovered it on the West: and that there was a small Light left on the lower part of the Sun that appeared like a Starr. And from Cranbrook in Kent, we are informed, by the Relation of the curious William Tempest Esq; R. S. S. that he observed there the Sun to be extinguished but for a Moment, and instantly to emerge again: So that the Limit past exactly over this Town, which is about 38 Geographical Miles from London, and very near the right Angle where the Perpendicular from London falls on the Line of the Limit, being 3'. 00" of Time to the Eastwards of London in the Latitude of 51°. 6', as near as I can gather.

How it past over *Suffex* we have not so authentick Relations, but have learnt that it was Total at Wadhurst beyond Tunbridge-wells, as also for some short time at Lewis; but that it was not so at Brightling, which Place being situated on an Eminence that has a commanding Prospect, all the Country to the Northward was seen in Darkness, whilst they there had some Benefit of a small Remainder of the Sun.

From these Observations we may conclude that this Limit came upon the Coast of *England*, about the middle between *Newhaven* and *Brightelmston* in *Suffex*, and passing by

(257)

by *Cranbrook* and *Bocton* left *Canterbury* about 4 Miles on the Right-hand, and quitted the Coast of *Kent* not far from *Hern* toward the antient *Regulbium*, now called *Reculver*. So that it seems scarce one third part of *Kent*, and not so much of *Suffex*, out of all the South Coast of *Great-Britain*, escaped being involved in this Darkness.

The Northern Limit, having past over a much greater Space, has had more Observers, and is not less curiously determined than the other. We find by the Account given by the Reverend Mr. Roger Prosser, Rector of *Haverford-West*, that the Eclipse was total there a Minute and half, whence it follows that *Haverford* was but about 6 Miles within the Shade, and therefore that it entred on *Pembrokeshire* about the middle of *St. Brides Bay*, leaving *St. David's* and *Cardigan* on the left Hand: and having traversed those two Counties and *Montgomeryshire*, it entred on *Shropshire*, leaving the Town of *Shrewsbury* 1'. 40" in the Shadow, as was observed there by Dr. *Hollings*: whereby it appears that *Shrewsbury* was about 8 Miles within the Limit. Thence it proceeded by the East-side of *Cheshire*, leaving *Whitchurch* and *Nantwich* a very little without, and passing by *Congleton* went over the Peak of *Darbyshire* into *Yorkshire*, and crost the great Northern Road between *Pontefract* and *Doncaster*, somewhat nearer the former than the latter. For by the Observations of that curious Gentleman *Theophilus Shelton Esq;* at *Darrington* about two Miles on this side *Pontefract*, (in Lat. 53°. 40' and Long. West from London 4'. 40") of time, as may be concluded from *Normood's Measure of a Degree* the Sun at 9^h. 1'. was reduced almost to a Point, which both in Colour and Size resembled the Planet *Mars*; but whilst he watched for the Total Eclipse, that Point grew bigger and the Darkness diminished; whence he argued the Limit to have been very little more Southerly. And since he has been informed that it was just Total in *Barnsdale*, three Miles South from thence.

And

Fig. 2 Pages 256 and 257 of *Philosophical Transactions of the Royal Society* (1715), containing Halley's compilation of eye-witness accounts from the vicinity of the southern and northern limits of the total solar eclipse of 1715.

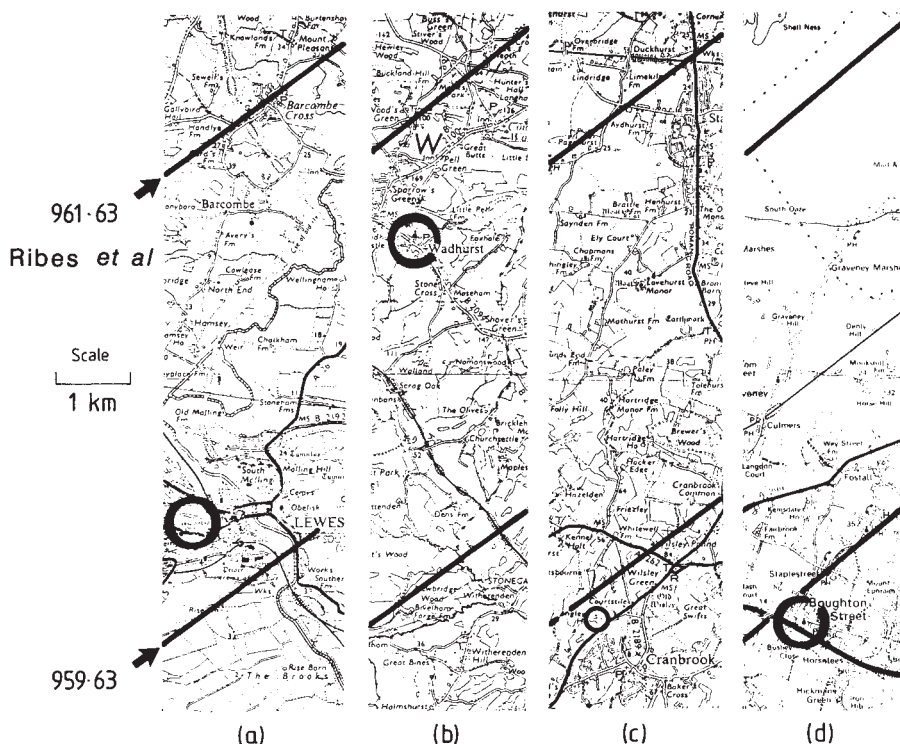


Fig. 3 The calculated southern limits of the path of totality of the 1715 eclipse in relation to Lewes (a), Wadhurst (b), Cranbrook (c) and Bocton (Boughton Street) (d). The limit using the current value of the Sun's semi-diameter is labelled 959.63 arc s. The limit using the early eighteenth century result of Ribes *et al.* is labelled 961.63.

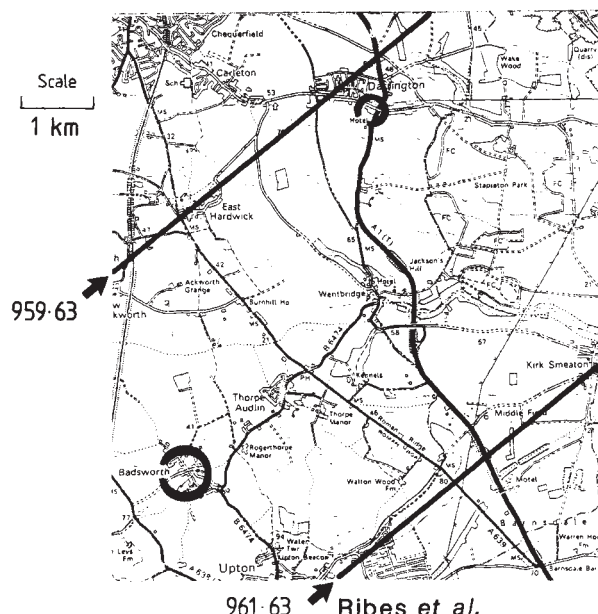


Fig. 4 The calculated northern limits of the path of totality of the 1715 eclipse in relation to Darrington and Badsworth. The limits have the same significance as in Fig. 3.

underestimated the effects of the optical imperfections on the image size in the telescopes used by Picard and La Hire, as argued recently by O'Dell and Van Helden⁹.

We thank David Calvert of the Royal Greenwich Observatory for the photographic reproduction of pages 256 and 257 of the *Philosophical Transactions of The Royal Society* for 1715. We are grateful to the Ordnance Survey for their kind permission to reproduce sections of the 1:50 000 maps shown in Figs 3 and 4.

Received 31 July; accepted 4 December 1987.

1. Ribes, E., Ribes, J. C. & Bartholot, R. *Nature* **326**, 52–55 (1987).
2. Archives Obs. Paris manuscripts D1. 14–16 (1666–1682) and D2. 1–10 (1683–1719).
3. Parkinson, J. H., Morrison, L. V. & Stephenson, F. R. *Nature* **288**, 548–551 (1980).
4. Parkinson, J. H. *Nature* **304**, 518–520 (1983).
5. Halley, E. *Phil. Trans. R. Soc.* **29**, 245–262; 314–316 (1715).
6. Dunham, D. W., Sofia, S., Fiala, A. D., Herald, D. & Muller, P. M. *Science* **210**, 1243–1245 (1980).
7. Watts, C. B. *Astr. Pap. Am. Eph. Naut. Al.* XVII (1963).
8. Stephenson, F. R. & Morrison, L. V. *Phil. Trans. R. Soc.* **A313**, 47–70 (1984).
9. O'Dell, C. R. & Van Helden, A. *Nature* **330**, 629–631 (1987).

Entrainment of trace-metal-enriched Atlantic-shelf water in the inflow to the Mediterranean Sea

Alexander van Geen, Paula Rosener & Edward Boyle

Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Surface waters of the Mediterranean Sea have higher trace-metal concentrations than open Atlantic surface waters¹. This observation could be explained by sources within the basin, but recent data from the Alboran Sea² indicate that trace metals may already be enriched in Atlantic waters flowing into the Mediterranean through the Strait of Gibraltar. Here we present new data documenting the presence of trace-metal-enriched source waters west of the Strait, and present a mathematical analysis showing that Cu, Ni, Cd and Zn distributions in the Alboran Sea can be explained as linear mixtures of four well-defined sources: (1) Spanish coastal

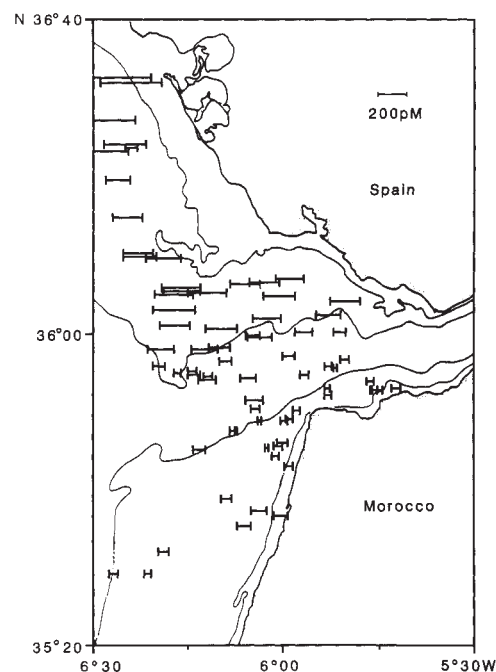


Fig. 1 Surface Cd concentrations during a survey (March–April 1986) of the eastern Atlantic and the Strait of Gibraltar. Station locations are centred on horizontal bars. Length of bar is proportional to Cd concentration. Bathymetry contours at 30 and 200 m.

waters, (2) North Atlantic Central Water, (3) Atlantic surface water, and (4) Mediterranean Levantine Intermediate Water. In this analysis, salinity and trace metals are used for the first time to deconvolve mixing between water masses of different origins.

Chemical budgets for the Mediterranean are influenced by vigorous exchange of Atlantic surface water for deep Mediterranean water through the Strait of Gibraltar. Fluxes through the Strait are 30 times higher than fresh-water inflow to the Mediterranean³. Nevertheless, surface Cu, Ni, Cd and Zn concentrations in the Mediterranean are higher than nutrient-depleted Atlantic surface water by up to 2, 2, 9 and 40 times respectively (refs 1, 2 and R. M. Sherrell and E.B., unpublished). If typical Atlantic surface waters were representative of the inflow, this difference would have to be maintained by inputs within the Mediterranean basin. As trace-metal enrichments are present in the inflowing plume², sources within the basin may not account for all of the enrichment. Here we show that trace-metal fluxes through the Strait of Gibraltar are enhanced by contributions from two trace-metal-enriched Atlantic sources: (1) Spanish shelf water and (2) subsurface open ocean water. Calculated average Cu, Cd and Zn concentrations of the inflow resemble the actual metal composition of Mediterranean waters for these metals. Therefore the net input of Cu, Cd and Zn within the basin may be dominated by external rather than internal sources.

There are two metal-rich water masses to the west of the Strait of Gibraltar. Previous work^{2,4} has shown that subsurface Atlantic waters in that region are enriched in Cd, Ni and Zn. Data from other parts of the ocean also show this deep enrichment, which is attributed to the decomposition of metal-rich biological debris at depth. Here we identify a second enriched water mass by collecting many surface samples from the eastern Atlantic on board USNS *Lynch* (26 March–19 April 1986) and on RV *Oceanus* cruise 176 (14–16 April 1986); Fig. 1 shows surface sampling locations. Samples were collected with a contamination-free underway pumping apparatus, pre-concentrated in the laboratory by modified Co-APDC (cobalt-ammonium pyrrolidinedithiocarbamate) co-precipitation⁵, and analysed by graphite furnace flameless atomic absorption. Cd is enriched in Atlantic surface waters on the Spanish shelf outside the Strait

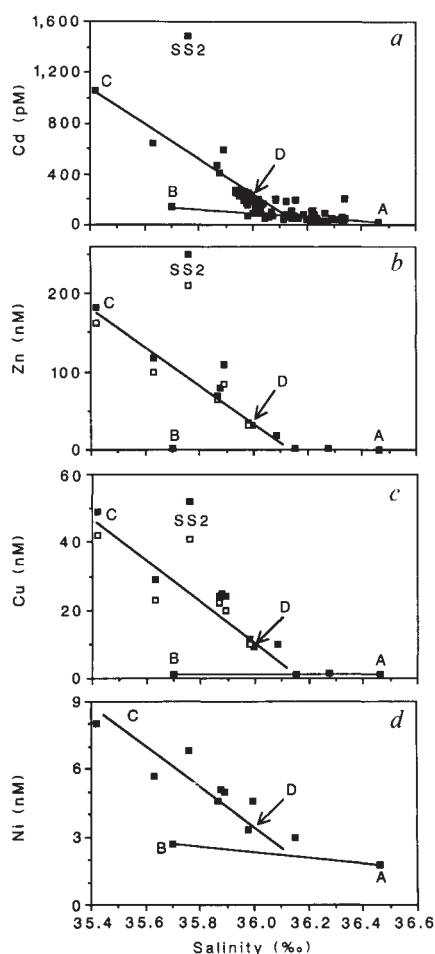


Fig. 2 Metal-salinity correlation in surface samples from Fig. 1. ■, Unfiltered samples; □, filtered through 0.4 μm Nuclepore filter. Atlantic profile end-members are indicated by A (surface) and B (400 m deep). By definition, samples resulting from conservative mixing of these two end-members fall on line segment AB. Positive deviations from this relationship which follow line segment CD are due to a shelf-specific trace-metal enrichment process. Arrow at D indicates metal composition chosen for representative shelf end-member. SS2 deviates from shelf-salinity line owing to variability in these waters. Consistent inter-element ratios argue against contamination.

of Gibraltar (Fig. 1). Cd concentrations of open-ocean surface Atlantic waters are usually <10 pM, whereas Cd concentrations of 200 pM are seen in waters overlying the Spanish shelf. These high-Cd waters are found as close as 30 km west of the Strait. Even a small contribution of waters from this area would significantly influence the composition of inflow to the Mediterranean Sea. Similar enrichments have been observed in other coastal regions^{6,7}.

Differences in the metal-salinity relationships of the two enriched water masses allow one to distinguish them (Fig. 2). Salinity and cadmium concentration gradients within the top 400 m of an open Atlantic profile taken 100 km west of the Strait can be described as a simple mixture of surface Atlantic and deeper (less saline) North Atlantic Central Water (Fig. 2a; data in ref. 2). Cd concentrations in low-salinity Spanish shelf samples follow a different metal-salinity relationship. Some Cd-enriched surface samples were also analysed for Cu, Ni and Zn. Cu and Zn are highly enriched in shelf waters, again showing clear differences between shelf and open-ocean metal-salinity relationships (Fig. 2b and c). Ni enrichment also occurs in shelf waters but is not as dramatic as for the other metals (Fig. 2d). Table 1 gives the composition of end-members representative of the different source areas.

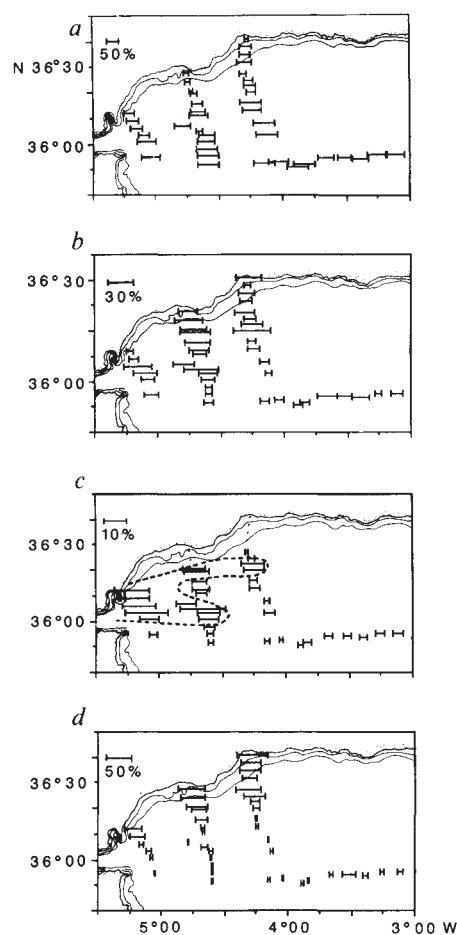


Fig. 3 End-member fractions in surface samples of the Alboran Sea east of the Strait of Gibraltar: a, surface Atlantic; b, deep Atlantic; c, shelf Atlantic; d, deep Mediterranean. Salinity and trace-metal data from refs 2 and R. M. Sherrell and E.B. (unpublished). Station locations are centred on horizontal bars. Length of bar is proportional to each end-member contribution. Bars are not drawn to the same scale from one figure to the other. The 8% contour is shown in Fig. 3c for shelf end-member fraction (broken line).

Metal-salinity systematics in shelf samples do not reveal trace-metal enrichment mechanisms because turbulent mixing over the shelf homogenizes these waters and obscures sources or processes leading to the enrichment. Several explanations for these enrichments can be considered. (1) A linear metal-salinity relationship could be due to mixing of metal-enriched river water and open-ocean sea water. A river source of this magnitude is not plausible here because extrapolation of the metal-salinity relationship to zero salinity requires an unrealistically high concentration of metals in the river water. (2) Local fallout of aeolian particles could enrich coastal waters more than open-ocean waters owing to the proximity of a continental aerosol source. But most studies suggest a more gradual decrease in aeolian deposition rates with distance from the coast than we observe for trace-metal enrichments in the Gulf of Cadiz. (3) Diagenesis within shelf sediments may result in diffusion of trace metals into the overlying water⁸. Such a source may be important, but it is difficult to establish the dominance of this mechanism compared with other plausible enrichment processes. (4) A coastal estuarine 'nutrient trap' can increase the residence time of biologically reactive elements such as trace metals over the shelf, thereby increasing their concentration in these waters. The 'source' in this case is simply the inflowing

Table 1 Tracer composition of model end-members and average inflow

Tracer	Surf. Atlantic	Deep Atlantic	Shelf Atlantic	Deep Mediterranean	Model error	Average inflow
Salinity (‰)	36.46	35.70	36.00	38.45	0.05	36.20
Cu (nM)	1.3	1.3	10.0	2.0	0.3	2.8
Ni (nM)	1.8	2.7	3.5	4.2	0.5	2.3
Cd (pM)	20	140	260	60	10	86
Zn (nM)	0.8	1.5	40.00	5.0	0.5	6.8

Following the notation in the text, the weighted least-square solution of end-member fractions for each surface sample is $f = (A' \cdot W \cdot A)^{-1} \cdot A' \cdot W \cdot d$ (refs 14 and 15, p.54). W is a diagonal matrix composed of the inverse of estimated model uncertainty for each tracer. Matrix W renders the problem non-dimensional by dividing all tracer values by corresponding model uncertainties. These are also listed above for each tracer. Uncertainty estimates are greater than (salinity) or equal to (trace metals) analytical errors and represent the degree to which the model is required to fit the data for a given tracer. The Kuhn-Tucker theorem provides an iterative solution to this problem as described in ref. 15, pp 126–131. Average inflow composition is based on contributions from surface, deep and shelf Atlantic end-members of 60, 25 and 15% respectively.

open-ocean sea water, from which metals and nutrients accumulate during particulate transfer and remineralization⁹. The circulation regime on the Spanish coast is conducive to such a nutrient trap because of significant freshwater input, whereas the Moroccan coast is not.

Although the evidence cannot distinguish unequivocally between these mechanisms, the sedimentary fluxes and nutrient trap hypotheses are the most plausible explanations for trace-metal enrichments on the Spanish shelf. Whatever the actual shelf enrichment process is, for our purpose the data can be interpreted as the result of mixing of a surface ((A) in Fig. 2) and a deep (B) Atlantic end-member as well as a shelf (D) end-member mixing in various proportions. A salinity of 36.00‰ was chosen for the shelf end-member based on the predominance of this water mass over the shelf west of the Strait. Trace-metal concentrations corresponding to this salinity are included in Table 1.

Four sources of water influence the trace-metal composition of the Alboran Sea inflow plume: (1) Atlantic surface waters, with low metal concentrations, (2) subsurface Atlantic waters; this trace-metal-rich water may be brought to the surface by mixing or upwelling during transport into the Alboran Sea, (3) Spanish continental shelf water, and (4) saline Levantine Intermediate water, which is entrained by mixing and upwelling. Owing to upwelling and clockwise rotation of the Alboran Sea gyre, this deeper Mediterranean water outcrops at the northern edge. Table 1 gives the composition of the deep Mediterranean water mass, which is nearly uniform below 200 m for these tracers², along with the three Atlantic end-members defined above. Aeolian particle transport and deposition also may cause surface-water metal enrichments in some environments, but we will demonstrate that Alboran Sea metal distributions can be accounted for without invoking atmospheric sources. Given the rapid circulation of surface waters in the region, conservative mixing (in terms of salinity and trace metal content) is assumed for the Alboran Sea. Residence times of ²¹⁰Pb in surface waters range from two years in the open ocean¹⁰ to less than a month in the highly productive Gulf of California¹¹. The assumption of conservative behaviour for elements that are less particle-reactive than Pb in the Alboran Sea therefore seems justified in view of the vigorous circulation. Based on this hypothesis, trace-metal distributions in the Atlantic inflow east of the Strait, which have been presented elsewhere², will be related to the inflow sources.

Assuming conservative mixing of salinity and trace metals, each Alboran Sea surface sample can be described as a linear combination of four distinct end-members. In this expression, the weight of an end-member is the fraction of the total sample contributed by each source. In matrix notation, this relationship can be expressed as $A \cdot f = d$, where A is a model matrix whose four columns contain the tracer composition of each end-member, d is the observed composition of a specific Alboran Sea sample and f is a column vector whose four elements are

the fractions of each contributing source which we want to estimate. The fractions of each end-member are constrained to lie in the range 0 to 1 and sum to unity. The system is over-determined, so analytical errors (and a small degree of non-conservative behaviour) mean that the solution will not match the data perfectly. Figure 3 shows the least-square solutions for f (see Table 1 for details) for each sample. Although each tracer constrains the final solution for all components, the solution for particular components is most sensitive to certain tracers: salinity dominates the estimate for the fraction of deep Mediterranean water, Zn and Cu dominate the estimate for the shelf end member, and Cd dominates the estimate for the deep Atlantic contribution. Model trace-metal concentrations match data for all five tracers within one 'model error' (defined in Table 1) at each of the 43 surface stations except three, where Cu and Ni residuals are significant. The very good agreement between model fit and observed tracer concentrations indicates that the conservative mixing model is consistent with the data. Even though temporal variations of trace metal relationships over the Atlantic shelf and in the Alboran Sea must still be determined, these results indicate consistency of two data sets collected four years apart during different seasons (June 1982 and April 1986).

Surface circulation features of the Alboran Sea (such as the clockwise rotation of the gyre and injection of Atlantic water through the Strait), are well characterized by satellite thermal imagery and can be compared with the areal distribution of end-members derived above. Surface Atlantic water clearly dominates the upper water column of the Alboran Sea (50–85% of total in Fig. 3a) with the exception of the northern margin of the gyre, where outcropping deep Mediterranean water contributes up to 60% to surface waters (Fig. 3d). This feature can be detected on the basis of salinity alone. Perhaps more interesting is the fact that Cd concentrations in this high-salinity region constrain the solution to include up to 30% of Atlantic deep water (Fig. 3b). A second feature, which is uniquely deduced from trace metal concentrations, is the entrainment of Spanish shelf water in the Atlantic inflow as visualized by the 8% contribution contour of this end-member in Fig. 3c. This result is mainly constrained by Cu and Zn data.

Because the composition of the Atlantic inflow in the Alboran Sea can be described as the result of mixing of three end-members identified in the eastern Atlantic, one can estimate the integrated contribution of each source to the Mediterranean at the time of the observations. Geostrophic calculations show that mass transport through a portion of the transect (from 36°01' N, 5°05' W to 36°09' N, 5°11' W) of the *¿Donde Va?* survey¹² is representative of the flux passing through the Strait. In addition, trace-metal and salinity data from a profile centred on this transect (ref. 2 and unpublished results) indicate that the relative composition of the combined Atlantic end-members is constant at a given location even though the total Atlantic contribution decreases with depth relative to that of the Mediterranean component. At each station, a homogenized Atlantic end-member

mixes vertically with Mediterranean water, but the composition of the pre-mixed Atlantic end-member varies from one surface station to the next, as Fig. 3 shows.

Based on these observations and a geostrophic velocity section from ref. 12, a representative composition of the total inflow can be calculated by weighting average metal fluxes of Atlantic origin vertically under each surface station and laterally along the geostrophic section. The resulting average composition of the inflow (see Table 1) is roughly equivalent to that of the deep Mediterranean except for Ni. Keeping in mind that the composition and volume¹³ of the inflow probably varies seasonally, it appears that a net source of trace metals within the Mediterranean basin exists for Ni but is less important for Cu, Cd and Zn.

We thank Dr Tom Kinder and Dr Nan Bray for providing ship space and time during the Gibraltar experiment on board USNS *Lynch*; and A. Spivack, S. Chapnick, D. Lea, W. Moore and E. Callahan for helping to collect the large number of samples. This research was sponsored by the Office of Naval Research.

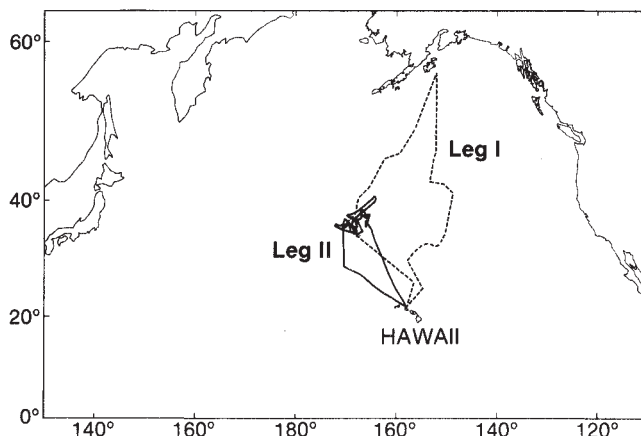


Fig. 1 Cruise track of the RV *Moana Wave* during the SEAREX North Pacific Experiment. Leg I, 28 April to 3 June 1986; leg II, 9 June to 14 July 1986.

Received 4 September; accepted 20 November 1987.

1. Spivack, A. J., Huested, S. S. & Boyle, E. A. in *Trace Metals in Seawater*, 505–512 (Plenum, New York, 1983).
2. Boyle, E. A., Chapnick, S. D., Bai, X. X. & Spivack, A. J. *Earth planet. Sci. Lett.* **74**, 405–419 (1985).
3. Tchernia, P. in *Descriptive Regional Oceanography* 142–147 (Pergamon, Oxford, 1980).
4. Statham, P. J., Burton, J. D. & Hydes, D. J. *Nature* **313**, 565–567 (1985).
5. Boyle, E. A., Huested, S. S. & Jones, S. J. *geophys. Res.* **86**, 8048–8066 (1981).
6. Bruland, K. W. & Franks, R. P. in *Trace Metals in Seawater*, 395–414 (Plenum, New York, 1983).
7. Kremling, K. *Nature* **303**, 225–227 (1984).
8. Heggie, D., Klinkhammer, G. & Cullen, D. *Geochim. cosmochim. Acta* **51**, 1059–1070 (1987).
9. Redfield, A. C., Ketchum, B. H. & Richards, F. A. in *The Sea* Vol. 5, 61–63 (Wiley, New York, 1963).
10. Bacon, M. P., Spencer, D. W. & Brewer, P. G. *Earth planet. Sci. Lett.* **32**, 277–296 (1976).
11. Bruland, K. W., Koide, M. & Goldberg, E. D. *J. geophys. Res.* **79**, 3083–3086 (1974).
12. Parilla, G. & Kinder, T. H. in *Proc. NATO Advanced Res. Workshop Oceanogr. Mediterranean, La Spezia*, 1983 (in the press).
13. Bormans, M., Garrett, C. & Thompson, K. *Oceanol. Acta* **9**, 403–414 (1986).
14. Mackas, D. L., Denman, K. L. & Bennett, A. F. *J. geophys. Res.* **92**, 2907–2918 (1987).
15. Menke, W. *Geophysical Data Analysis - Discrete Inverse Theory* (Academic, Orlando, 1984).

Evidence for $\geq C_3$ alkyl nitrates in rural and remote atmospheres

Elliot Atlas

Department of Oceanography, Texas A&M University, College Station, Texas 77843-3146, USA

Interactions of odd-nitrogen compounds with non-methane hydrocarbons and a variety of radical species have been shown to have a potential impact on atmospheric chemistry on local, regional, and even global scales^{1–9}. To understand the interactions of atmospheric nitrogen compounds requires an accurate budget of individual reactive nitrogen species. But recent attempts to obtain a mass balance between total reactive nitrogen (NO_y) and individual reactive nitrogen species ($NO + NO_2 + HNO_3 + PAN$ (peroxyacetyl nitrate) + PPN (peroxypropionyl nitrate) + particulate NO_3^-) have not accounted for all NO_y ^{9,10}. One class of compound proposed to account for some of the NO_y deficit is the alkyl nitrates ($RONO_2$)^{5,10–12}, although ambient measurements have not been available to test this hypothesis. Here I report measurements of $RONO_2$ made in the North Pacific atmosphere, providing the first direct evidence for existence and transport of $\geq C_3$ alkyl nitrates in the troposphere. I also observed a variation in alkyl nitrate concentration which suggests these compounds may be used as unique tracers of 'polluted' air masses to remote areas. These observations are significant because relatively long-lived alkyl nitrates can serve as reservoirs of reactive nitrogen in the

remote troposphere, where they can be converted to NO_x , a key catalyst in ozone formation and destruction^{3,4}.

Alkyl nitrates were measured in air samples collected during a SEAREX (Sea-Air Exchange) cruise in the North Pacific Ocean during April–July 1986 (see Fig. 1). During the cruise two types of samples were taken. Long-term (16–20 h) samples were obtained from a specially constructed air-sampling tower ~7 m above the deck at the bow of the ship. Typical sample volume for these samples was ~300 l. Short-term samples were taken from a mast extending several feet in front of the bow. These samples were usually taken during the morning for approximately 30 min. These samples were ~12 l. All sampling was conducted only with wind coming over the bow of the ship at >5 m.p.h. This excluded possible contamination from ship exhaust. Continuous measurement of condensation nuclei and other chemical data further confirmed the absence of ship contamination during sampling.

Samples were obtained by drawing air at 200–400 ml min⁻¹ through two 5-mg charcoal traps in series. For all analyses reported here there was $<5\%$ breakthrough of analytes into the backup trap. After sample collection, the traps were spiked with internal standard and extracted three times with purified benzene. High-resolution capillary gas-chromatographic analysis of the extract was performed using a Varian 3500 GC with an electron-capture detector (ECD). Several major unknown peaks were always observed in the sample extracts during the cruise. Subsequent tests revealed these unknowns as alkyl nitrates.

Identification of the alkyl nitrates was performed on preserved samples in a shore-based laboratory by using gas chromatography-mass spectrometry in the negative-ion chemical ionization (NICI) mode. A mass spectrum of the major unknown compounds from marine samples (Fig. 2) showed that the unknown compound was also present in relatively high concentration in local ambient air; thus most mass spectrometry experiments were done on local air extracts to preserve the few marine samples available. In the local samples not only were there major fragments at m/e 46 and 71, but the m/e 46 fragment was present in a homologous series of compounds following $m/e = 57 + n \times 14$, with $n = 0–5$ (see Fig. 3); the single most abundant compound was at $m/e = 71$. The $m/e = 46$ fragment suggested NO_2 , and $\Delta(m/e) = 14$ suggested CH_2 , but the negative ion fragment at $m/e = 71$ was a puzzle.

The conclusion that alkyl nitrates could produce the observed NICI mass spectrum was suggested by reports that, in the CI process, alkoxide ions are stabilized by loss of H_2 ¹⁴. This phenomenon is widely observed in ion-cyclotron resonance studies of reactions of alkoxide ions with many organic species¹⁵. NICI-MS of purchased and/or synthesized alkyl nitrates

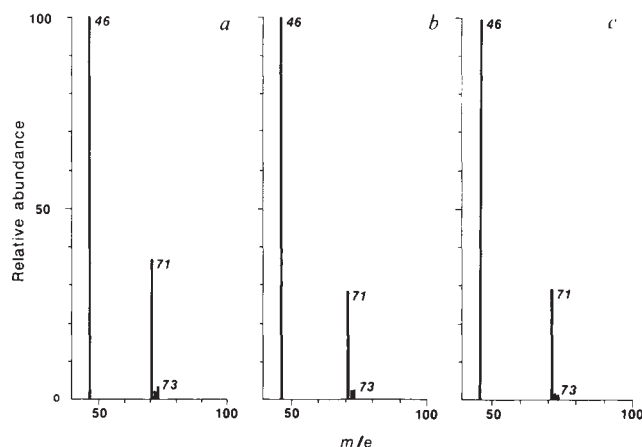


Fig. 2 Negative-ion chemical ionization mass spectra of authentic standard of 2-butyl nitrate (a) and spectra obtained from atmospheric samples from College Station, Texas (b), and from the North Pacific (c).

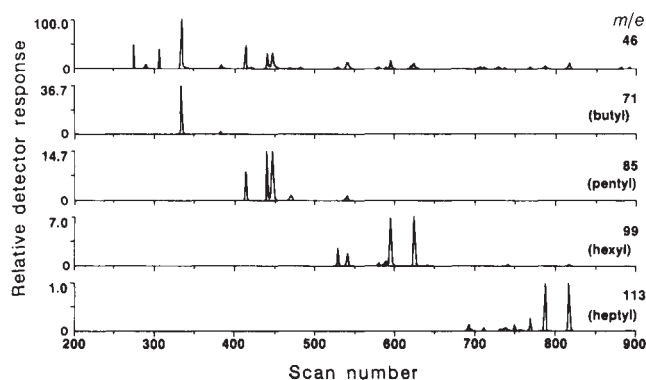


Fig. 3 Extracted ion profiles (negative ion) of air sample from College Station, Texas. The peak of $m/e = 46$ is NO_2 . Major peaks in the chromatogram represent mostly secondary alkyl nitrates of a given alkyl chain length.

verified this process under the conditions described here. The additional H_2 loss from the butoxide fragment ($m/e = 73$) produces the observed $m/e = 71$ fragment of butyl nitrate (see Fig. 2). Thus the series of compounds observed in the local air samples was identified as C_3 – C_7 alkyl nitrates.

GC–MS verification with authentic standards showed that the main alkyl nitrates in the marine samples were 2-butyl, 2-pentyl and 3-pentyl nitrates. Butyl nitrate was approximately four times more abundant than the pentyl nitrates in these samples. Smaller amounts of propyl nitrate, *n*-butyl nitrate and *n*-pentyl nitrate also were suggested by the MS data, but interferences in the ECD did not allow measurement of these minor components. One additional pentyl nitrate isomer was found but it is not identified and its concentration is not included in the results discussed below; GC retention data and theoretical considerations suggest that it is 2-isopentyl nitrate. (2-methyl-1-propyl-nitrate may also be present in the samples and would coelute with 2-butyl nitrate, but this compound is expected only in small amounts relative to 2-butyl nitrate.) The observed distribution of secondary and primary alkyl nitrates is in agreement with theoretical and laboratory experiments^{16–18}. Secondary alkyl nitrates are the major nitrate products predicted from OH-initiated oxidation of hydrocarbon and subsequent reaction with NO^{16} .

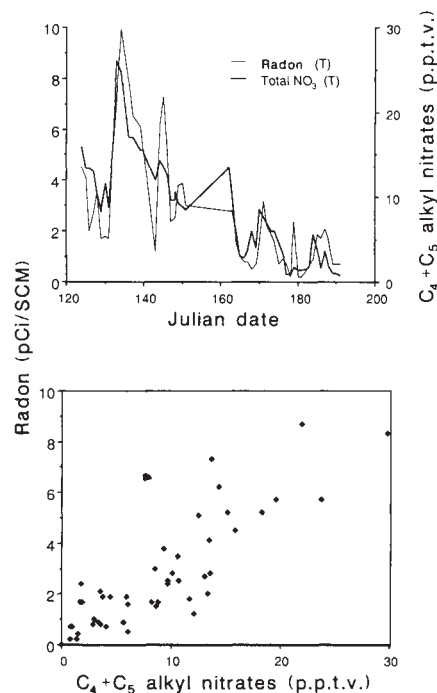


Fig. 4 Correlation of $\text{C}_4 + \text{C}_5$ alkyl nitrates with radon during the North Pacific cruise. Alkyl nitrate data shown here are from long-term samples; radon is average radon measured several times over the nitrate sample period.

The concentration of alkyl nitrates measured in this study seem to be a significant fraction of the reactive nitrogen budget in the marine troposphere. The total concentration of 2-butyl and 2- and 3-pentyl nitrates in all samples ranged from 0.8 to 37 p.p.t.v. The mean alkyl nitrate concentration in the long-term samples was 9 ± 7 p.p.t.v., whereas the mean of the short-term samples was 2 p.p.t.v. higher. It should be emphasized, too, that alkyl nitrates measured here do not represent the total alkyl nitrate concentration in the marine troposphere. For example, not all pentyl nitrates were measured, C_3 nitrates were identified and not measured, and it has been suggested that C_1 and C_2 nitrates may also be present⁹. The maximum concentrations of alkyl nitrates found here approach the levels of PAN measured over the Pacific Ocean (~ 50 p.p.t.)¹⁹ and may amount to several per cent of NO_x (~ 200 – 500 p.p.t.)^{7,9,13,20}.

During the cruise I observed variations in alkyl nitrates that suggested long-range transport. For example, Fig. 4 compares the concentration and temporal variation of $\text{C}_4 + \text{C}_5$ alkyl nitrates with those of atmospheric radon. Increases in radon concentration indicate rapid transport of continental air to the study site during leg I of the cruise. This transport is further substantiated by air-parcel trajectory analysis²¹ and observations of radionuclides from the Cherbonyl accident²² during the same period.

Given normal background concentrations of NO_x (~ 10 – 100 p.p.t.) and alkanes (10 – 100 p.p.t.) in mid-ocean regions^{7,14,23,24}, and the expected mechanism of RONO_2 formation in environments with high organic and NO_x concentrations, it is likely that the major source of nitrates to the study area was long-range transport from the Asian continent. The excellent correlation of RONO_2 and radon (Fig. 4) supports this hypothesis. Although I favour the idea of long-range transport as the course of most RONO_2 to the mid-Pacific, *in situ* formation of RONO_2 in the marine troposphere has been suggested (J. Roberts, personal communication) and cannot be excluded as another possible source of RONO_2 for these samples.

Also, because RONO_2 compounds are easily formed, we must consider the possibility that sampling artefacts contribute to observed concentrations. Several lines of evidence, however, suggest that artefact formation is not a major problem for these samples. First, alkyl nitrates are measured in comparable concentrations in samples varying over an order of magnitude in sample volume. If artefact formation were a problem, for example by oxidation of hydrocarbons adsorbed on the sampler, then considerably higher concentrations should be observed in larger volume samples. This is not the case. Second, the variation in RONO_2 concentration correlates with radon. To obtain the observed correlation with radon and an order-of-magnitude variation in RONO_2 concentration solely as a result of artefact formation, one must hypothesize rapid transport or oceanic emission of RONO_2 precursors to the mid-Pacific without prior photochemical reactions. However, it is more likely that RONO_2 will form from precursors in a high- NO_x (urban) environment or during a 5–10-day transport time rather than during a 30-minute sampling period. However, I plan further testing of the sampling method to determine if artefact formation might bias the measurement.

There is now good evidence that alkyl nitrates are present in the remote marine troposphere and in rural continental areas. Mass spectral studies have confirmed the identity of a series of RONO_2 ($R \geq C_3$) compounds in the ambient atmosphere. These measurements are consistent with model predictions¹¹ and with laboratory studies^{18,25}. The mode of formation of RONO_2 compounds in high NO_x environments and the relative stability of these compounds relative to OH reaction¹⁸ suggest that RONO_2 may serve as a storage mechanism of reactive nitrogen in the troposphere much the same way as PAN¹⁹. Furthermore, my measurements in remote regions indicate that alkyl nitrates may reach concentrations that could affect ozone formation chemistry in the marine troposphere. If this is so, more data on RONO_2 are required to assess OH- NO_x -HC chemistry in the atmosphere.

I thank the Oceanography and Atmospheric Sciences Division of NSF for support (OCE-84-15726) as part of the SEAREX programme. I thank H. Y. Tong and S. Cross for assistance in air sampling; the officers and crew of the RV *Moana Wave* for their cooperation and assistance during the cruise; Mr J. Lindgren and associates at the Texas Air Control Board for use of NICI-MS facilities; and Dr J. Roberts and R. Dickerson for helpful suggestions and comments on the manuscript. The University of Rhode Island Center for Atmospheric Chemistry Studies provided the radon data and air trajectory analysis.

Received 26 June; accepted 9 November 1987.

- Altshuller, A. P. *Atmos. Environ.* **20**, 245–268 (1986).
- Crutzen, P. J. *A. Rev. Earth planet. Sci.* **443**–472 (1979).
- Crutzen, P. J. *Pure appl. Geophys.* **106**–108, 1385–1399 (1973).
- Chameides, W. & Walker, J. C. G. *J. geophys. Res.* **78**, 8751–8760 (1973).
- Dickerson, R. R. *J. geophys. Res.* **90**, 10,739–10,743 (1985).
- Kasting, J. F. & Singh, H. B. *J. geophys. Res.* **91**, 13,239–13,256 (1986).
- Logan, J. A. *J. geophys. Res.* **88**, 10,785–10,807 (1983).
- Logan, J. A., Prather, M. J., Wofsy, S. C. & McElroy, M. B. *J. geophys. Res.* **86**, 7210–7245 (1981).
- Singh, H. B. *Environ. Sci. Technol.* **21**, 320–327 (1987).
- Fahey, D. W. *et al. J. geophys. Res.* **91**, 9781–9793 (1986).
- Calvert, J. & Madronich, S. *J. geophys. Res.* **92**, 2211–2220 (1987).
- Grosjean, D. *Environ. Sci. Technol.* **17**, 13–17 (1983).
- Dickerson, R. R. *Atmos. Environ.* **18**, 2585–2593 (1984).
- Hayes, R. N., Sneldron, J. C., Bowie, J. H. & Lewis, D. E. *Aust. J. Chem.* **38**, 1197–1208 (1985).
- Noest, A. J. & Nibbering, N. M. M. *Adv. Mass. Spectrom.* **8A**, 227–237 (1979).
- Atkinson, R. A. & Lloyd, A. C. *J. phys. Chem. Ref. Data* **13**, 315–444 (1984).
- Pitts, J. N. *Jr J. phys. Chem.* **80**, 1948–1950 (1976).
- Atkinson, R., Aschmann, S. M., Carter, W. P. L., Winer, A. M. & Pitts, J. N. *Jr Int. J. chem. Kinet.* **14**, 781–788 (1982).
- Singh, H. B. & Salas, L. J. *Nature* **302**, 326–328 (1983).
- Noxon, J. F. *J. geophys. Res.* **88**, 11,017–11,021 (1983).
- Merrill, J. T. *Searex News.* **10**, 2–6 (1987).
- Uematsu, M. & Duce, R. A. *Maritimes* **30**(3), 1–4 (1986).
- Ridley, B. A., Carroll, M. A. & Gregory, G. L. *J. geophys. Res.* **92**, 2025–2047 (1987).
- Rudolph, J. & Khedim, A. *Int. J. environ. analyt. Chem.* **20**, 265–282 (1985).
- Carter, W. P. L., Lloyd, A. C., Sprung, J. L. & Pitts, J. N. *Jr Int. J. chem. Kinet.* **11**, 45–101 (1979).

Historical decline of red spruce populations and climatic warming

Steven P. Hamburg* & Charles V. Cogbill†‡

* University of Kansas, Lawrence, Kansas 66045, USA

† Center for Northern Studies, Wolcott, Vermont 05680, USA

Within the scientific as well as the popular press there is a growing interest in the declining growth rates and increased mortality of red spruce (*Picea rubens* Sarg.) in the eastern United States^{1,2}, and the possibility of a causal relationship with anthropogenic pollutants, particularly acidic deposition^{3–7}. To evaluate this relationship we need to consider the long-term (200 year) population trend of this species (which has not been done to date). Potentially long-term changes are being viewed with a short-term perspective, possibly leading to the identification of spurious causal relationships. Using historical and current compositional data from old-growth stands in the spruce-hardwood zone we constructed a 180-year trend in forest composition for central New Hampshire. These data indicate a major decrease in the size of red spruce populations from 1800 to the present. It is proposed that the major driving force of this decline has been a warming trend in both mean annual and mean summer temperatures.

Tree cores have been the basis of most studies of forest growth decline to date, yet such data do not provide insights into long-term population trends. Sampling of an individual tree's past history using tree cores precludes important information about stand development or population patterns, or requires the assumption of long-term stability of species composition.

The long generation time of most forest trees requires that the interval necessary for examining all but the most radical changes in forest populations is several times greater than the life span of most researchers. No single researcher can effectively view population changes without relying on historical material, yet few such data exist in North America. The need for multiple historical sources to establish a quantitative record of population changes has hindered the establishment of a long-term perspective.

In a small area of central New Hampshire near the Hubbard Brook Experimental Forest, where high levels of acid precipitation are well documented⁸, historical data are available on pre-disturbance forest composition of mid-elevational forests from independent sources: witness trees used by early surveyors to demarcate lot and town boundaries and forest inventories. When combined with present forest composition data from undisturbed forests in the same area, shifts in species composition over the past 200 years may be examined with confidence. Within a 15 km radius of Welch Mountain, Waterville, NH there are witness tree data from two towns (Campton and Woodstock) surveyed in the 1770s to 1790s⁹, species composition of one hundred and twenty 0.4-ha plots collected in 1903 prior to intensive logging¹⁰, and the Bowl Research Natural Area, a 206 ha all-aged undisturbed (old-growth) forest^{11–13} (Fig. 1). As altitudinal gradients are present within the forest communities of the region¹⁴, data on species composition at a single elevational band within the spruce-hardwood zone¹⁵ (550–710 m) were extracted from all of these sources.

Since the modern population data are from a single site the potential for site specific variation suggests the need for information from another old-growth forest. Such a stand exists 50 km to the northeast near Mountain Pond in Chatham, New Hampshire. This stand is near additional 1903 survey plots in the Low and Burbank Grant, and when combined these data provide an independent corroboration of the Waterville-Bowl temporal pattern.

The witness tree data (Table 1) from Woodstock and Campton

‡ Present address: Goddard College, Plainfield, Vermont 05667, USA.

Table 1 Relative densities of dominant species in the virgin spruce-hardwood forests of central New Hampshire from 1770 to 1984, and percentage of individuals in each taxon with a diameter-at-breast-height greater than 20 cm

Location	Witness trees		Direct measurement				
	Campton	Woodstock	Waterville	Low and Burbank	Bowl Research Natural Area		Mountain Pond
Date of measurement	1770-1797	1794	1903	1903	1973	1974	1984
Logging	none	none	very limited cutting if any	very limited cutting if any	none	none	none
Elevation range (m)	580-700	580-700	580-700	unknown	580-640	580-700	520-550
Number of trees	29	40	5,420	506	384	74	77
Species	Forest composition(%)						
<i>Picea rubens</i>	45	38	29	42	6	7	5
<i>Fagus grandifolia</i>	21	32	35	24	38	47	49
<i>Betula</i> sp.	10	22	24	20	25	31	8
<i>Acer</i> sp.	3	0	8	6	30	14	31
<i>Tsuga canadensis</i>	7	0	3	5	0	0	0
<i>Abies balsamea</i>	7	8	<1	0	1	0	0
Other	6	0	<1	3	0	1	7

are consistent, providing an excellent indication of the pre-settlement forests of central New Hampshire. Since five surveys over a 20-year period surveyed Campton, there is a low probability of systematic surveyor bias^{16,17}. The 1903 plot data are more quantitative than the witness tree data, representing the average of fifty 0.4-ha plots, yet might include some plots cut-over for spruce^{10,18,19}.

There is a clear and consistent pattern (Table 1) of species composition between the 1903 plot data and witness tree data, with red spruce sharing a roughly equal role with beech, with lesser amounts of birch, maple, hemlock and fir. The cause of the slightly lower percentage spruce in one group of the 1903 data (that is, 29% at Waterville versus 45% and 38% Campton and Woodstock respectively) is unknown; it could represent the effects of limited cutting, differences in data collection, site differences and/or the beginning of a broad decline. The species composition of the pre-settlement forest of the region fits well within Braun's¹⁵ characterization of a spruce-hardwood forest.

The modern forest composition of the old-growth Bowl

Research Natural Area, when compared with the pre-settlement data, presents striking differences (Table 1). The paucity of red spruce 5-7% in the modern old-growth forest and the greater importance of maple and birch indicate major differences between the modern and pre-settlement data. The age distribution of red spruce within the Bowl Research Natural Area also indicates a declining population²⁰.

The uncut stand at Mountain Pond is at 520-550 m elevation, very similar to the sites near Welch Mountain, and was inventoried by Carboneau in 1984 (ref. 21 and personal communication). Mountain Pond is 25 km from the Low and Burbank Grant which was the site of additional survey plots in 1903. The contrast of 1903 and 1984 data from these more northerly stands showed a pattern of spruce decline that is consistent with that seen in the Waterville data (Table 1). Shifts in forest species composition clearly have occurred over the past 100 to 200 years in central New Hampshire; concomitant with a dramatic decrease in red spruce has been an increase in maple and birch. For over 100 years observations have been made of unexplained death of red spruce in old-growth forests of the northeastern United States further supporting these data²²⁻²⁴.

From the data presented in Table 1 little can be said about timing of the spruce decline. However, evidence of recent spruce mortality (16 spruce snags (trees lacking tops) ha⁻¹ at Mt Pond)²¹ in the modern old-growth forests accounts for less than 25% of the spruce loss since 1800. Thus, initiation of the decline has to predate observations of growth decline since 1960². Though there were high levels of spruce present in 1903, the factors affecting the subsequent decline may have begun to impact the population prior to that date, particularly with respect to regeneration. Spruce regeneration (trees 10 to 20 cm diameter-at-breast-height, with breast height taken as 1.4 m present in the 1903 Waterville plots (69 trees ha⁻¹) and Low and Burbank plots (57 trees ha⁻¹)¹⁰ are not present in the modern old-growth populations (7 trees ha⁻¹ at the Bowl and 9 trees ha⁻¹ at Mt Pond)^{10,21}.

Over the scale of millennia the composition of the spruce-hardwood forests have not been constant, reflecting changing climate and the advance and retreat of species^{25,26}. Red spruce was virtually absent from the lower slope (<1,100 m) forests in northern New England prior to 2,000 yr BP^{27,28}. Its appearance coincides with a return of spruce pollen in many southern sediment cores²⁹⁻³². This shift is usually attributed to climatic cooling after the most recent post-glacial warm period when spruce retreated, 'presumably to higher elevations'²⁹ or into refuges on restricted wet or rocky habitats at intermediate elevations (C.V.C. and P. S. White, unpublished data). However, at 650 m elevation in the White Mountains of New Hampshire,

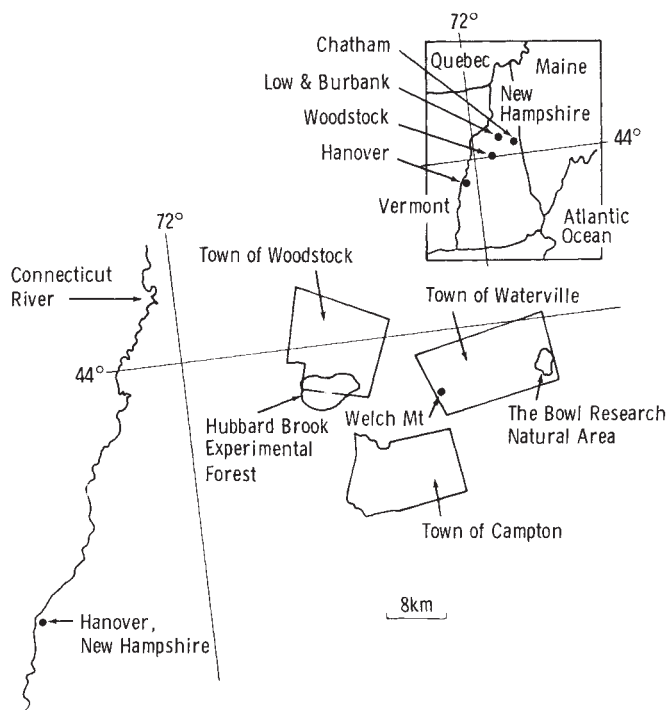


Fig. 1 Map of Central New Hampshire showing the location of study areas.

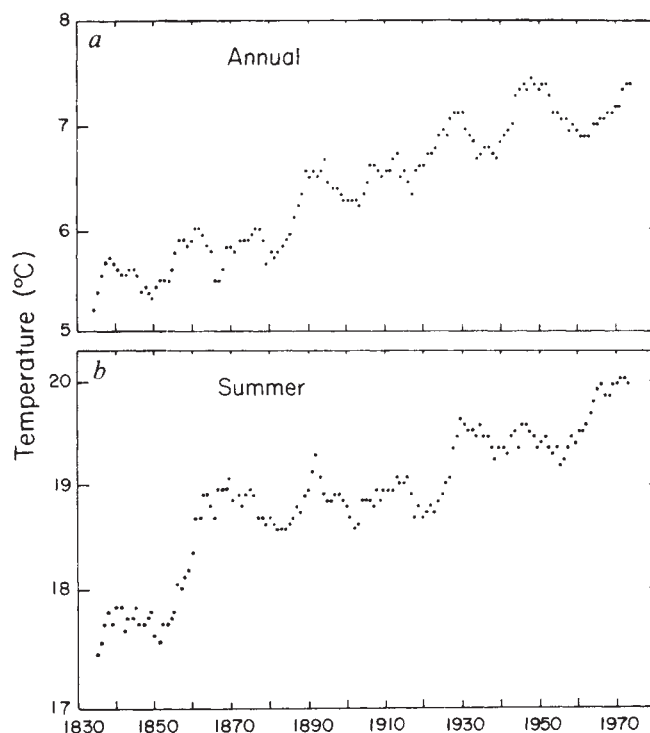


Fig. 2 Running average of mean temperatures at Hanover, New Hampshire, for the 10-yr period following the year indicated. Data were derived from daily records in the Meteorological Journals in the Dartmouth College Archives and in Climatological Data of NOAA (Ashville, NC). The averages are the mean of daily maximum and minimum temperatures for 1870 to 1983. The thrice daily readings from 1835 to 1869 were normalized to the maximum-minimum period using 15 years of observations using both methods. Summer includes the months of June, July and August. Occasional missing monthly data were estimated using the long-term Amherst, MA record³⁸.

spruce pollen and macrofossil deposits show relatively constant values from 2000 yr BP until 200 yr BP²⁵⁻²⁸. Beginning at the time of settlement by Europeans (*circa* 1780), there has been a substantial decrease in spruce pollen deposited in the region. Clearing of the land and selective logging of spruce by settlers have contributed to this decrease, but recent changes in spruce abundance are also consistent with decline at its lower limits after the cold period (Little Ice Age) which ended in the early 1800s. Interpretation of the palynological record is hampered by the lack of temporal resolution over the period of both climatic and human induced change.

Spurr first suggested that recent warming of the annual temperature could be causing spruce decline³³. The clear indication of an increased spruce presence in the palynological record at the lower limits of its range during climatic cooling indicates a negative response of this species to warming conditions²⁹⁻³². The coniferous/deciduous boundary is closely correlated with summer temperatures (C.V.C. and P. S. White, unpublished data) and spruce radial growth is consistently correlated (negatively) with growing season temperatures of the previous year³⁴⁻³⁶. A computer simulation of the forest composition in northern New England indicates the replacement of red spruce by hardwood with a 2 °C warming during the growing season³⁷. The climate record from central New England (Hanover, NH) shows warming of both yearly (1.7 °C) and summer (2.2 °C) temperature over the past 150 years (Fig. 2) and comparable to that seen at Amherst, MA³⁸. Elimination of spruce from the spruce-hardwood zone (200 m wide) on the slopes of northern New England, where this species is recent and marginal, would require just this magnitude of temperature change. For example, the increase in summer temperature at Hanover, over the past 150 years is the equivalent to over 400 m upward displacement

of the coniferous/deciduous boundary²⁷. It is unclear if the deciduous/coniferous boundary is the result of conditions of regeneration or competition among established canopy trees. In either case the decline would necessarily have a long lag-time (100-200 years in computer simulation of abrupt climate change³⁷) as climate changes slowly and individual old trees remain without replacement for their lifespan of several hundred years. This evidence suggests that spruce population decline in the spruce-hardwood zone over the past 100 years is driven, in large part, by increased average temperature. It is important to note that spruce decline is not restricted to the lower elevation spruce-hardwood population, and that there is evidence of greater recent decline in higher elevation populations³. The relative role of temperature and/or pollution in the decline of the higher elevation spruce populations cannot be addressed by the data presented.

Without long-term paired plots it is difficult to separate environmentally induced trends from those resulting from land use practices, chronic pollution and/or a pathogen. However, we do hypothesize that the decline of spruce populations in the mid-elevation forests of central New Hampshire has been initiated by a warming climate. Currently we may be witnessing a natural population decline possibly aggravated by pollution or pathogen factors.

An important consideration in viewing spruce decline is that it is a long-term phenomenon probably involving a host of environmental factors. The dramatic short-term declines currently evident are real, but in looking for causal relationships they must be considered in the context of the long-term changes demonstrable with historical data.

We thank F. H. Bormann, L. Carbonneau, L. Conkey, M. B. Davis, G. Geballe, L. Huenneke, J. Hull, A. H. Johnson, C. W. Martin, T. G. Siccama, D. Wang and W. Winner for helpful discussions and reviews of an earlier manuscript. We also thank L. Carbonneau, L. Conkey, R. Bradley, C. W. Martin, New Hampshire Historical Society, National Archives, Town of Campton, and Town of Woodstock for providing data.

Received 30 June; accepted 26 November 1987.

- Siccama, T. G., Bliss, M. & Vogelmann, H. W. *Bull. Torrey Bot. Club* **109**, 162-168 (1982).
- Hornbeck, J. W. & Smith, R. B. *Can. J. For. Res.* **15**, 1199-1201 (1985).
- Johnson, A. H. & McLaughlin, S. B. in *Report of the Committee on Monitoring and Yields in Acidic Deposition*, 200-230 (National Research Council, National Academy Press, Washington, DC, 1986).
- Shabecoff, P. *N.Y. Times* (2/26/1984) p. 1, 33.
- Johnson, A. H. & Siccama, T. G. *Tappi J.* **67**, 68-72 (1984).
- Johnson, A. H. & Siccama, T. G. *Environ. Sci. Technol.* **17**, 294A-305A (1983).
- Johnson, R. H., Turner, R. S. & Lord, D. G. in *Direct and Indirect Effects of Acid Deposition on Vegetation* Vol. 5, (ed. Linthurst, R. A.) (Butterworth, Boston, 1984).
- Likens, G. E., Bormann, F. H., Pierce, R. S., Eaton, J. S. & Johnson, N. M. *Biogeochemistry of a Forested Ecosystem* (Springer, New York, 1977).
- Hamburg, S. P. thesis, Yale Univ. (1984).
- Chittenden, A. K. in *State of New Hampshire Biennial Report of the Forestry Commission 1903-04*, 1-219 (Concord, New Hampshire, 1904).
- Leak, W. B. *U.S.D.A. For. Serv. Res. Note NE-186* (1974).
- Leak, W. B. *U.S.D.A. For. Serv. Res. Note NE-181* (1973).
- Martin, C. W. *U.S.D.A. For. Serv. Res. Note NE-244* (1977).
- Bormann, F. H., Siccama, T. G., Likens, G. E. & Whittaker, R. H. *Ecol. Monogr.* **40**, 373-388 (1970).
- Braun, E. L. *Deciduous Forests of Eastern North America* (Free Press, New York, 1950).
- Bourdo, E. A. Jr *Ecology* **37**, 754-768 (1956).
- Siccama, T. G. *Amer. Mid. Nat.* **85**, 153-172 (1971).
- Weigle, W. C. thesis, Yale Univ. (1904).
- Chapman, H. H. *The Collected Papers of Herman H. Chapman* Vol. 117 (Yale University Manuscript Collection, New Haven, 1906).
- Leak, W. B. *Ecology* **56**, 1451-1454 (1975).
- Carbonneau, L. thesis, Univ. New Hampshire (1986).
- Weiss, M. J., McCreery, L. R., Millers, I., Miller-Weeks, M. & O'Brien, J. T. *U.S.D.A. For. Serv. Northeastern Area NA-TP-11* (1985).
- Fox, W. F. in *Annual Report of the Forestry Commission for 1894* (Forestry Commission of the State of New York, Albany 1895).
- Mook, P. V. & Eno, H. G. *U.S.D.A. For. Serv. For. Res. Notes* **59**, 1-2 (1956).
- Davis, M. B. in *The Quaternary of the United States* (eds Wright, H. E. Jr & Frey, D. G.) 377-401 (Princeton University Press, 1965).
- Terasmae, J. & Anderson, T. W. *Can. J. Earth Sci.* **7**, 406-413 (1970).
- Davis, M. B., Spear, R. W. & Shane, L. C. K. *Quat. Res.* **14**, 240-250 (1980).
- Spear, R. W. thesis, Univ. Minnesota (1982).
- Likens, G. E. & Davis, M. B. *Verh. Internat. Verein. Limnol.* **19**, 982-993 (1975).
- Whitehead, D. R. *Quat. Res.* **12**, 333-357 (1979).
- Davis, R. B., Bradstet, T. E., Stuckenrath, R. Jr & Borns, H. W. Jr. *Quat. Res.* **5**, 435-465 (1975).
- Davis, M. B. in *Late-Quaternary Environments of the United States, Vol. 2, The Holocene* (ed. Wright, H. E. Jr.) 166-181 (University of Minnesota Press, 1984).

33. Spurr, S. H. *Amier. J. Sci.* **251**, 682-688 (1953).
 34. Conkey, L. thesis, Univ. Arizona (1982).
 35. Cook, E. R., Johnson, A. H. & Blasing, T. J. *Tree Phys.* **3**, 27-40 (1987).
 36. McLaughlin, S. B., Downing, D. J., Blasing, T. J., Cook, E. R. & Adam, H. S. *Oecologia* **72**, 487-501 (1987).
 37. Davis, M. B. & Botkin, D. B. *Quat. Res.* **23**, 327-340 (1985).
 38. Bradley, R., Ives, P. T. & Eischeid, J. *The Climate of Amherst* (University of Massachusetts, Amherst, 1987).

Role of surface speciation in the low-temperature dissolution of minerals

Alex Blum & Antonio Lasaga

Department of Geology and Geophysics, Yale University,
 PO Box 6666, New Haven, Connecticut 06511, USA

Surface species control the mechanism and rate of low-temperature silicate dissolution and precipitation reactions in dilute solutions. This has been repeatedly emphasized¹⁻⁷. The interaction between dissolved species and silicate surfaces involves exchange and adsorption phenomena, which equilibrate rapidly, and are accessible to determination by surface titration^{8,9}. Here we report a comparison of surface species concentrations with the dissolution rates of olivine and albite which indicates that while the dissolution rates are a complex function of solution pH, the dissolution rates have a simple first-order dependence on surface concentrations of specific surface species. Thus, the interpretation of the kinetic data for olivine and albite dissolution becomes simple when the chemical speciation at the mineral surface is incorporated, and this approach holds great promise for unifying the rate data for many important silicates.

The kinetics of rock-water interactions are dominated by chemical reactions occurring at the solid-solution interface. However, most low-temperature experiments have concentrated on the effect of the bulk solution composition on mineral growth and dissolution^{5-7,10-16}. This type of data provides only limited and indirect evidence of the chemical reactions actually occurring at the mineral surface¹⁷. Surface analytical techniques such as X-ray photoelectron spectroscopy¹⁸⁻²⁰, Auger spectroscopy²¹ and hydrogen profiling²² provide valuable quantitative information about the overall composition of the mineral surface. The analysis must, however, be performed at high vacuum, and

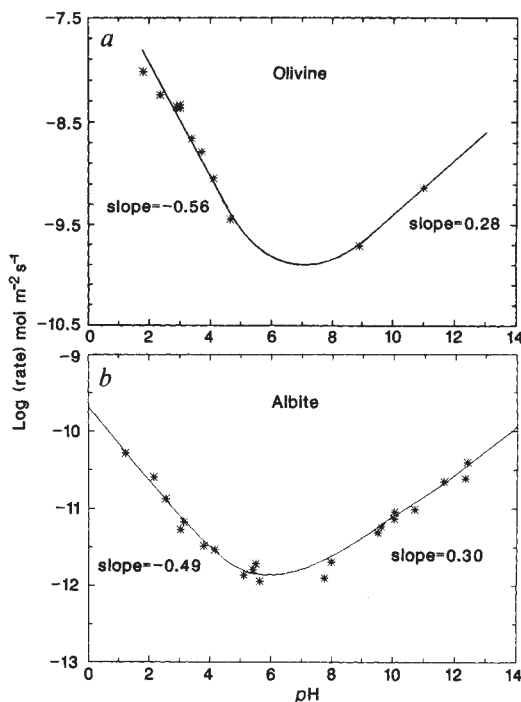


Fig. 1 Showing log of the dissolution rate ($\text{mol m}^{-2} \text{s}^{-1}$) versus pH of a, olivine, and b, albite (data from Chou and Wollast¹²).

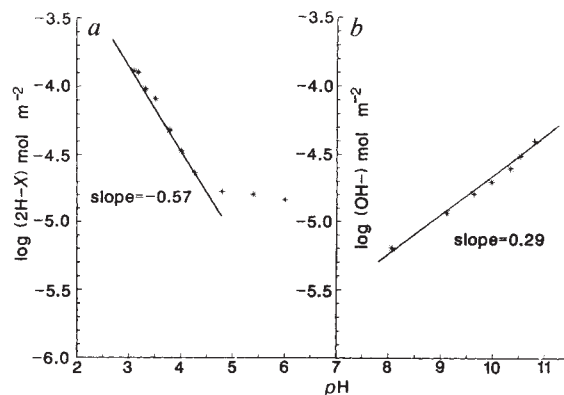


Fig. 2 Olivine surface concentration of 2H-X and S-OH^- (mol m^{-2}) versus pH.

cannot yield other important information about the interaction between the surface and solution.

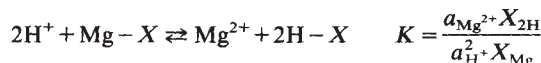
The technique of surface titration presents a fast, simple and powerful method for determining the chemical speciation on mineral surfaces^{8,9}. Finely ground Balsam Gap olivine (Fo_{93}) and Amelia albite (Ab_{97}) were suspended in water and titrated with acid and base while simultaneously monitoring the pH and Mg or Na concentration with ion-selective electrodes. Chemical analysis of the solution after the titration showed that the total amount of Si and/or Al released during these titrations was completely negligible. The total amount of H^+ adsorbed on the surface was calculated by subtraction of a titration blank. The net charge on the surface is the difference between the total amount of H^+ adsorbed and the quantity of charged cations released into solution. Olivine dissolution rates were determined using a flow-through fluidized bed reactor. Amelia albite dissolution rates were determined by Chou and Wollast using a similar apparatus^{5,11,12}. The steady-state dissolution of both minerals was stoichiometric and no secondary phases were precipitated. Surface areas were determined by N_2 and Kr adsorption using the BET adsorption isotherm.

Figure 1 shows the experimentally determined dissolution rates of olivine and albite as a function of pH. The dissolution rate is determined from the release rate of Si into solution. The dependence of the dissolution rate on pH has been in recent years expressed as $\text{Rate} \propto a_{\text{H}^+}^{n+1-7,11-13,15,25}$, where a_{H^+} is the H^+ activity in solution, and n is an experimentally derived value. The rate data for olivine and albite yield the following rate laws:

Olivine	Albite	
$\text{Rate} = ka_{\text{H}^+}^{0.56}$	$\text{Rate} = ka_{\text{H}^+}^{0.49}$	$2 \leq \text{pH} \leq 6$
$\text{Rate} = ka_{\text{H}^+}^{0.28}$	$\text{Rate} = ka_{\text{H}^+}^{0.30}$	$8 \leq \text{pH} \leq 12$

The dissolution rates of albite and olivine both show the same qualitative dependence on pH, namely a decrease of rate with pH in the acid region, and an increase of rate in the basic region. This behaviour is typical of silicate minerals¹⁶. Several theoretical attempts have been forthcoming to interpret this type of rate data^{1-3,6,23,24}. However, the variable and fractional values for n , typical of most silicates are very difficult to interpret mechanistically, and have presented a major stumbling block to understanding mineral dissolution kinetics. Some workers¹⁻⁴ have suggested that the H^+ adsorption isotherm may be involved.

The titration of the olivine surface and analysis of the solutions indicates that below pH 5 there is a stoichiometric and completely reversible exchange of 2H^+ for Mg^{2+} . This is a simple ion-exchange reaction in which two H^+ are removed from solution and replace Mg^{2+} in a surface site, releasing Mg^{2+} into solution. This exchange is stoichiometric, and results in no net change in surface charge. An equilibrium constant can be calculated for this exchange reaction:



where X_{2H} and X_{Mg} are the fraction of surface exchange sites occupied by $2H^+$ and Mg^{2+} respectively, and the equilibrium constant K equals $10^{3.89}$. Above pH 8, the surface develops an increasingly negative surface charge ($S-OH^-$).

If we plot the log surface concentrations of these species, $\log(c_{2HX})$ in the acid region and $\log(c_{S-OH^-})$ in the basic region versus pH (Fig. 2), and force a straight line through the steep portion of the isotherm, the results are in very close agreement with the rate data.

$$c_{2HX} \propto a_{H^+}^{0.59}$$

$$c_{S-OH^-} \propto a_{H^+}^{-0.29}$$

The agreement extends from acid to basic solutions, even though the exponents change drastically. The surface chemical speciation combined with the rate law suggests that (1) in the acid region, the dissolution mechanism depends upon the exchange of $2H^+$ for Mg^{2+} and can be described by the equation,

$$\text{Dissolution rate} = kc_{2HX}^{1.0}$$

where c_{2HX} is the surface concentration (mol m^{-2}) of exchange sites occupied by $2H^+$, and k is the reaction constant for this mechanism, which equals $10^{-4.5} \text{ s}^{-1}$; and, (2) in the basic region, olivine dissolution is dominated by a different mechanism involving deprotonation of the tetrahedral oxygens, and development of a negative surface charge. Dissolution in the basic region can be described by the equation,

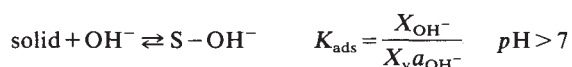
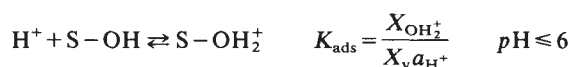
$$\text{Dissolution rate} = kc_{S-OH^-}^{1.0}$$

where c_{S-OH^-} is the surface concentration of negatively charged surface sites, and k is the reaction constant for this dissolution mechanism, which equals $10^{-4.5} \text{ s}^{-1}$.

In the neutral pH region, the concentration of both $2H-X$ and $S-OH^-$ surface species are so small that they cannot be accurately measured by surface titration. In addition, there is greater uncertainty in the dissolution rate, due to both the instability of the pH in near-neutral unbuffered solutions, and the possible increased role of crystal defects and impurities at very low dissolution rates. Consequently, it is uncertain whether the low dissolution rate of albite and olivine in the neutral region is controlled by low surface concentrations of $2H-X$ and $S-OH^-$, by dissolution at crystal defects, or by a separate dissolution mechanism which cannot be distinguished in these experiments.

The albite surface behaves rather differently as a function of pH than that of olivine. Na^+ on the albite surface was exchanged with H^+ in less than 1 minute. Thereafter, the Na^+ concentration in solution stayed nearly constant throughout the titration, and Na^+ was never readsorbed at any pH . This behaviour is in agreement with similar experimental data of Chou and Wollast¹². Because the Na^+-H^+ exchange sites are H^+ saturated over the entire pH range, it is not the Na^+ for H^+ exchange which is controlling the experimentally observed dependence of the albite dissolution rate on pH .

Figure 3 shows the surface concentration of H^+ adsorbed and desorbed from the albite surface in excess of the amount of Na^+ removed, as a function of pH . This H^+ adsorption/desorption results in a net change in the surface charge, and is probably the protonation/deprotonation of surface tetrahedral oxygens. These adsorption reactions can be described by Langmuir adsorption isotherms:



where a_{H^+} and a_{OH^-} are the activity of H^+ and OH^- in the bulk solution, $X_{OH_2^+}$ and X_{OH^-} are the mole fraction of surface exchange sites occupied by H^+ and OH^- , and X_v is the mole

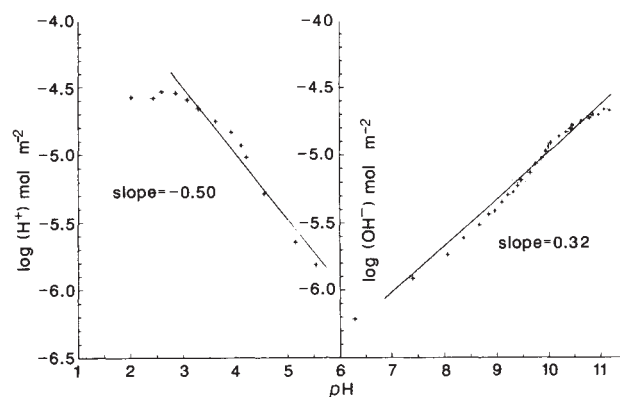


Fig. 3 Albite surface concentration of H^+ and OH^- (mol m^{-2}) versus pH .

fraction of vacant exchange sites. In the acid region, the equilibrium constant $K_{ads} = 10^{-3.90}$, and in the basic region $K_{ads} = 10^{-4.69}$.

Figure 3 plots the log of the surface concentrations of $S-OH_2^+$ and $S-OH^-$ as a function of pH . Note the flattening of the curve in the extreme acid region, which we feel reflects the near-saturation of the proton surface sites. Fitting, therefore, the steep portion of the acid adsorption isotherm and the basic isotherm to straight lines, slopes of -0.50 and 0.32 are obtained, in good agreement with the dependence of the dissolution rate on pH .

$$c_{S-OH_2^+} \propto a_{H^+}^{0.50} \quad pH < 6$$

$$c_{S-OH^-} \propto a_{H^+}^{-0.32} \quad pH > 7$$

This implies that the dissolution rate of albite is directly proportional to the surface concentrations of positive surface charge in the acid region, and negative surface charge in the basic region.

$$\text{Dissolution rate} = k_a c_{S-OH_2^+}^{1.0} \quad pH < 6$$

$$\text{Dissolution rate} = k_b c_{S-OH^-}^{1.0} \quad pH > 7$$

where $c_{S-OH_2^+}$ and c_{S-OH^-} are the surface concentrations of $S-OH_2^+$ and $S-OH^-$ respectively. The k s are reaction constants for the two mechanisms, $k_a = 10^{-6.5} \text{ s}^{-1}$ and $k_b = 10^{-6.1} \text{ s}^{-1}$. Once speciation is taken into account, the rate law and mechanism of the mineral-water reaction is elucidated. An exception seems to be the very low pH dissolution rate data for albite ($pH \leq 2$) which, although changing slopes slightly at pH 1 (Fig. 1b), does not flatten as much as the adsorption isotherm (Fig. 3a). More work is needed to sort out the details of the very low pH region in this case (this is not true for olivine).

Two aspects should be emphasized. First, these adsorption isotherms are not simple functions of the bulk solution compositions, and do not have the form $c_{ads} = ka_{H^+}^n$ over the entire pH range (see Fig. 3a, for example). The fractional values of n reported in the literature for the dependence of the dissolution rate on H^+ is a simplification as is also the straight line through the steep portion of the adsorption isotherm on a $\log(X_{ads})$ versus pH diagram. Second, the kinetics of these adsorption reactions are very fast, of the order of seconds to minutes, and completely reversible. These adsorption reactions do not in themselves present a kinetic barrier to olivine or albite dissolution. Note that the dissolution rate of BeO has a second-order dependence on the adsorbed protons, and Al_2O_3 a third-order dependence²⁵, in contrast to the first-order dependencies found for olivine and albite.

The measurement of surface speciation has several important implications. First, it elucidates the dissolution mechanism, allowing a much more rigorous treatment of dissolution and precipitation kinetics. Second, once the adsorption characteristics of a trace species on a specific surface is understood using

simple surface titrations, the need for long and tedious dissolution experiments is reduced to finding the rate constant and exponent for the surface species, rather than empirically determining the dissolution rate throughout the full range of solution compositions. However, we must caution that much more dissolution kinetic work must still be done to verify the cases where first-order dependence is valid, and we need to explore the surface chemistry of silicates much more thoroughly. Nonetheless, understanding the surface chemistry of dissolution promises to unify the rate data and dissolution mechanisms of a wide range of silicates.

Received 8 June; accepted 5 November 1987.

1. Stumm, W., Furrer, G., Wieland, E. & Zinder, B. in *The Chemistry of Weathering* (ed. Drever, J. I.) 55–74 (Reidel, Dordrecht, 1985).
2. Helgeson, H. C., Murphy, W. M. & Aagard, P. *Geochim. cosmochim. Acta* **48**, 2405–2432 (1984).
3. Lasaga, A. C. in *Kinetics of Geochemical Processes* (eds Lasaga, A. C. & Kirkpatrick, R. J.) 35–48 (Mineral. Soc. Am., Washington DC, 1981).
4. Rimstidt, J. D. & Dove, P. M. *Geochim. cosmochim. Acta* **50**, 2509–2516 (1986).
5. Chou, L. & Wollast, R. *Am. J. Sci.* **285**, 963–993 (1985).
6. Holdren, G. R. & Speyer, P. M. *Am. J. Sci.* **285**, 994–1026 (1985).
7. Tole, M. P., Lasaga, A. C., Pantano, C. & White, W. B. *Geochim. cosmochim. Acta* **50**, 379–392 (1986).
8. Stumm, W. & Morgan, J. J. *Aquatic Chemistry* (Wiley, New York, 1981).
9. Haug, C. P. in *Adsorption of Inorganics at Solid–Liquid Interfaces* (eds Anderson, M. A. & Rubin, A. J.) 183–218 (Ann Arbor Sci., Ann Arbor, Michigan, 1981).
10. Berner, R. A. & Holdren, G. R. *Geochim. cosmochim. Acta* **43**, 1173–1186 (1979).
11. Chou, L. & Wollast, R. *Geochim. cosmochim. Acta* **48**, 2205–2217v (1984).
12. Chou, L. & Wollast, R. in *The Chemistry of Weathering* (ed. Drever, J. I.) 75–96 (Reidel, Dordrecht, 1985).
13. Holdren, G. R. & Speyer, P. M. *Geochim. cosmochim. Acta* **49**, 675–681 (1985).
14. Grandstaff, D. E. *Geochim. cosmochim. Acta* **42**, 1899–1901 (1978).
15. Grandstaff, D. E. in *Rates of Chemical Weathering* (eds Colman, S. M. & Dethier, D.) 41–59 (Academic, New York, 1986).
16. Lasaga, A. C. *J. geophys. Res.* **89**, 4009–4025 (1984).
17. Helgeson, H. C. *Nature* **325**, 667 (1987).
18. Petrovic, R., Berner, R. A. & Goldhaber, M. B. *Geochim. cosmochim. Acta* **40**, 537–548 (1976).
19. Fung, P. C., Bird, G. W., McIntyre, G. G., Sanipelli, G. G. & Lopata, V. J. *Nucl. Tech.* **51**, 188–196 (1980).
20. Schott, J. & Berner, R. A. *Geochim. cosmochim. Acta* **47**, 2233–2240 (1983).
21. Hochella, M. F., Ponader, H. B., Turner, A. M. & Harris, D. W. *Geochim. cosmochim. Acta* (in the press).
22. Petit, J., Mea, G. D., Dran, J., Schott, J. & Berner, R. A. *Nature* **325**, 705–707 (1987).
23. Helgeson, H. C. *Geochim. cosmochim. Acta* **35**, 421–469 (1971).
24. Helgeson, H. C. *Geochim. cosmochim. Acta* **36**, 1067–1070 (1972).
25. Furrer, G. & Stumm, W. *Geochim. cosmochim. Acta* **50**, 1847–1860 (1986).

Unusual Early Cretaceous birds from Spain

J. L. Sanz*, J. F. Bonaparte† & A. Lacasa‡

*Unidad de Paleontología, Dto. Biología, Facultad de Ciencias, Universidad Autónoma, Cantoblanco 28049, Madrid, Spain

†Museo Argentino de Ciencias Naturales, Avda Angel Gallardo 470, 1405 Buenos Aires, Argentina

‡Amics de la Paleontologia, Institut d'Estudis Ilerdencs, Apartat 79, Lleida, Spain

The Neocomian Spanish outcrops of Montsec (province of Lérida) and the new one of Las Hoyas (province of Cuenca) have yielded several avian remains in the last few years. Several isolated feathers have been reported from Montsec, and a specimen of some feathered wing bones has recently been found. Las Hoyas has yielded an isolated feather and a nearly articulated small fossil bird that lacks the skull. This new specimen, reported here, presents a combination of derived (strut-like coracoids, pygostyle) and primitive (pelvic girdle, sacrum, hind limb) character states. If one considers *Archaeopteryx*, Ornithurae and the new Spanish fossil bird, it seems clear that the latter taxon is the sister group of Ornithurae (extant birds and all other fossil birds that are closer to recent forms than is *Archaeopteryx*).

Three years ago a new Early Cretaceous outcrop was found in the province of Cuenca, east-central Spain. It was called 'Las Hoyas' ('The Dales'), and it is probably equivalent to the formation 'Calizas de la Huérguina'¹, Neocomian (probably

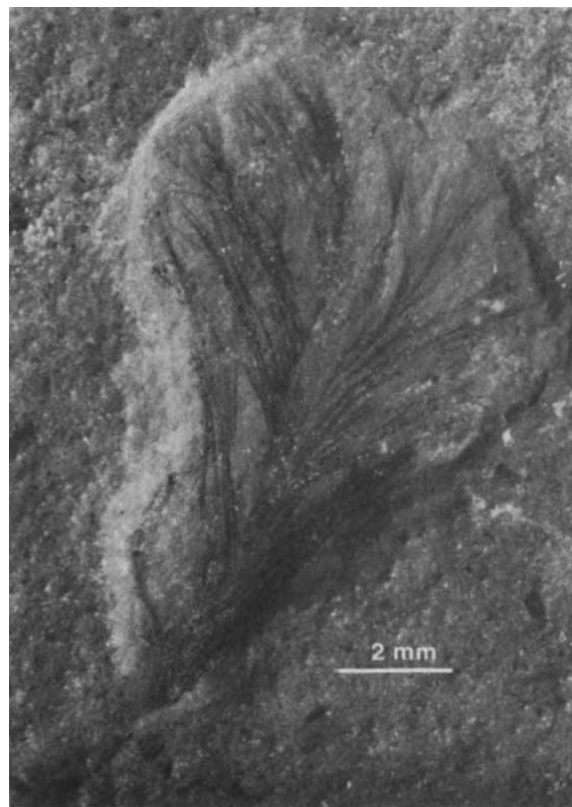


Fig. 1 Isolated feather from the Neocomian outcrop of Las Hoyas (Cuenca, Spain).

Hauterivian–Early Barremian) in age (some 130–120 Myr BP). The lithology is made of lithographic limestones from a lacustrine environment. Floral and faunal remains are relatively abundant, arthropods (insects and crustaceans) especially so. The vertebrate fauna is made up mainly of fishes along with an urodelan, an anuran, a chelonian and a little lizard. The faunal and floral assemblage from Las Hoyas seems, as a whole, very similar to that of the classic Neocomian outcrops of Montsec.

Both Neocomian Spanish outcrops, Montsec and Las Hoyas, have yielded several avian specimens in the last few years. Several feathers from the first one have been described². Some remains of feathered bones have recently been found from the Montsec. They are scattered in a lithographic limestone matrix. Both humeri, ulnae and radii can be seen, as well as a partially exposed furcula (wish-bone) with an acute hypocleidium. In *Archaeopteryx*, feathers have been preserved as impressions or even casts³. Unlike *Archaeopteryx*, the feathers of the Montsec specimen are dark carbon films.

To date Las Hoyas has yielded an isolated feather (Fig. 1) and an almost complete little fossil bird which was found by a local collector, Mr Armando Díaz Romeral who kindly loaned it to us for study. The specimen (number LH022R, Figs 2 and 3) is now housed in the Unidad de Paleontología of the Universidad Autónoma de Madrid. This Early Cretaceous Spanish bird is small (femur length 15 mm) and is exposed in a limestone slab mainly in lateral (left) view. The fossil is nearly articulated, but it lacks the skull and some cervical vertebrae, as well as the carpus and manus.

A striking characteristic of this Early Cretaceous bird is its combination of primitive (pelvic girdle, sacrum and hind limbs) and derived (coracoids and pygostyle) character states. These features allow it to be placed in an intermediate phylogenetic position between *Archaeopteryx* and later birds. The number of dorsal vertebrae is 11. This figure is intermediate between that of *Archaeopteryx* (13–14 segments⁴) and the greatly reduced



Fig. 2 The new Mesozoic bird (LH022R) from the Neocomian outcrop of Las Hoyas (Cuenca, Spain).

number in extant birds (4–6 segments⁵). The number of sacral vertebrae is probably five, which is symplesiomorphic with *Archaeopteryx* and all dinosaurs⁶. One of the most notable features of this small Spanish fossil bird is the appearance of a conspicuous pygostyle. This structure fuses a relatively large number of caudal vertebrae, probably 10–15, and seems larger than that of modern birds. Between the pygostyle and the sacrum are a series of eight free caudal vertebrae. The presence of a pygostyle, even in such an unusual form, appears to be a synapomorphy of this fossil bird with the Ornithurae (*sensu* Gauthier⁶ and Cracraft⁷). There is no evidence of gastralia, an apomorphic character state in relation with *Archaeopteryx* and theropod dinosaurs.

More 'evolutionary novelties' can be seen in the pectoral region. The coracoid is typically avian, strut-like, with a conspicuous head and an expanded distal end for the connection with the sternum. The furcula has a developed styloideus furcular process (hypocleidium) seen in anterior view that seems very similar to that of the feathered Montsec remains reported above.

The pelvic girdle is poorly preserved. The ilium is probably primitive, similar to that of *Archaeopteryx* and the coelurosaurian dinosaurs. The pre-acetabular portion seems to be relatively deep, and its surface greater than that of the post-acetabular region. The acetabulum is relatively small, similar to that of *Archaeopteryx*. The pubis is missing. The ischium lacks the peculiar outline of that of *Archaeopteryx*⁸, being more primitive than that of the Bavarian Jurassic bird. The pelvic elements are unfused⁷, a plesiomorphic character state shared with *Archaeopteryx*. The features of the hind limb are primitive too. The astragalus (which has a conspicuous theropodan ascending process) and calcaneum are free, not fused to the tibia. The free distal tarsals cap the proximal area of metatarsals, which are also unfused. Some synapomorphies of Carinatae, like the strut-like coracoid, can be found in a bird even more primitive than the Ornithurae. The Spanish Early Cretaceous bird lacks several of the synapomorphies proposed by Cracraft⁷ for the Ornithurae, like a complete fused tarsometatarsus or ossified uncinata processes.

The new fossil reported here represents a previously unknown level in the organization of birds, intermediate between *Archaeopteryx* and later birds. Considering this Early Cretaceous bird, *Archaeopteryx* and the Ornithurae, based on the character

states mentioned above, it seems clear that this new taxon is the sister group of the Ornithurae. The main synapomorphies which define this clade are the appearance of a pygostyle and the strut-like coracoids. The new fossil suggests that the early evolution of birds was firmly and rapidly influenced by the requirements of flight. So, the avian condition of scapular girdle and wing skeleton contrast with the conservative level of pelvic girdle, sacrum and hind limb. Even the appearance of a pygostyle precedes other typical avian features such as the fusion of the proximal tarsals to the tibia or the co-ossification of the metatarsals. The fossil also suggests a relatively rapid change in morphology from that of *Archaeopteryx* to that of Early Cretaceous birds, as was suggested by Kurochkin⁹, based on *Ambiortus*. It represents the only Spanish fossil material actually related to Mesozoic birds. Some isolated feathers from the Montsec have been denominated *Ilerdopteryx*². In 1902

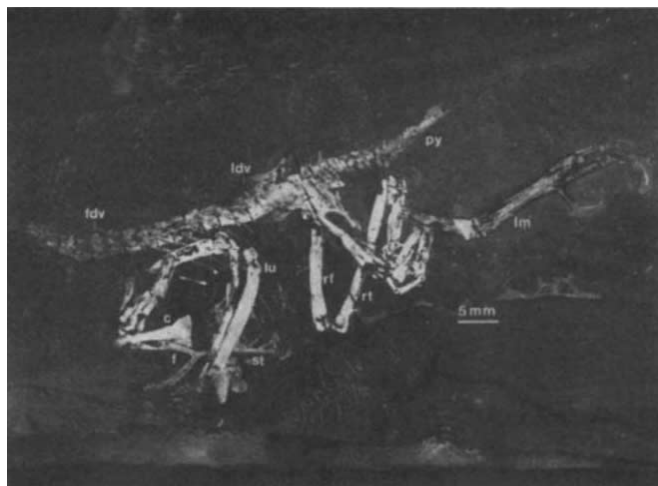


Fig. 3 Ultraviolet-induced fluorescence photograph of the new Spanish Neocomian bird (Las Hoyas, province of Cuenca). c, coracoid; f, furcula; fdv, first dorsal vertebra; ldv, last dorsal vertebra; lm, left metatarsals; lu, left ulna; py, pygostyle; rf, right femur; rt, right tibia; st, sternum.

Vidal¹⁰ reported a bird from Montsec lost by a workman. *Cose-saurus*^{11,12} is probably related to the prolacertid lepidosaurs¹³, and *Priscavolucris*¹⁴ was based on very doubtful remains.

We thank P. Wellnhofer, K. Padian, J. Ostrom, J. Cracraft and an anonymous reviewer for critically reading the manuscript and making valuable suggestions, and G. Rougier for the preparation of the specimen, which was made in the Museo Argentino de Ciencias Naturales (Buenos Aires). The photographs were taken by G. F. Kurtz. The excavations in Las Hoyas have been supported by funds from the 'Consejería de Educación y Cultura, Junta de Comunidades de Castilla-La Mancha'.

Received 2 July; accepted 23 November 1987.

1. Vilas, L. et al. in *El Cretácico de España* 457-513 (Univ. Complutense, Madrid, 1982).
2. Lacasa, A. *Ilerda* 46, 227-238 (1985).
3. Rietschel, S. in *The Beginnings of Birds* (eds Hecht, M. K., Ostrom, J., Viohl, G. & Wellnhofer, P.) 251-260 (Freunde des Jura-Museums Eichstätt, 1985).
4. Ostrom, J. *Biol. J. Linn. Soc. Lond.* 8, 91-182 (1976).
5. Van Tyne, J. & Berger, A. J. *Fundamentals of Ornithology* (Wiley, New York, London, Sidney & Toronto, 1976).
6. Gauthier, J. in *The origin of Birds and the evolution of Flight* (ed. Padian, K.) 1-98 (Mem. California Acad. Sci., 1986).
7. Cracraft, J. *Paleobiology* 12, 383-399 (1986).
8. Wellnhofer, P. in *The Beginnings of Birds* (eds Hecht, M. K., Ostrom, J., Viohl, G. & Wellnhofer, P.) 113-122 (Freunde des Jura-Museums Eichstätt, 1985).
9. Kurochkin, E. N. *Cretac. Res.* 6, 271-278 (1985).
10. Vidal, L. M. *Mems. R. Acad. Cienc. Artes Barcelona* 4, 263-274 (1902).
11. Ellemberger, P. & De Villalta, J. F. *Acta geol. hisp.* 9, 162-168 (1974).
12. Ellemberger, P. *Mem. Trav. E. P. H. E. Inst. Montpellier* 4, 91-117 (1978).
13. Sanz, J. L. & López-Martínez, N. *Geobios* 17, 747-753 (1984).
14. Gómez-Pallero, J. E. *Boln. Geol. Min.* 90, 333-346 (1979).

Polarization-opponent interneurons in the insect visual system

Thomas Labhart

Zoologisches Institut der Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Recent behavioural experiments dealing with the mechanism of polarized skylight navigation in bees indicated that processing of *e*-vector information in the visual system involves antagonistic interaction between polarization-sensitive photoreceptors^{1,2}. Here we report electrophysiological recordings from polarization-opponent interneurons in the optic lobe of crickets. These neurons receive antagonistic input from polarization sensitive photoreceptors with orthogonally arranged analyser orientations. Although polarization-sensitive interneurons have previously been reported from the visual system of crabs³ and goldfish^{4,5}, this is the first demonstration of polarization-opponent units.

In both bees and crickets polarization vision is mediated by a small group of specialized ommatidia situated at the dorsal rim of the compound eye^{6,7}. Each of these ommatidia contains two sets of homochromatic, strongly polarization-sensitive photoreceptors with orthogonally arranged analyser directions⁸⁻¹¹, as indicated by the orthogonal orientation of the microvilli (Fig. 1). The receptors have greatly enlarged visual fields due to optical specializations (bees, corneal pore canals; crickets, no retinal screening pigment)^{8,10-12}. These common features suggest that bees and crickets use similar mechanisms in *e*-vector detection and orientation.

This study was aimed towards an understanding of the physiology of polarization-sensitive interneurons within the optic lobe (medulla) of field crickets. The dorsal rim area was stimulated with flashes of polarized light with different *e*-vector directions or with a continuously rotating *e*-vector. The huge and strongly overlapping visual fields have the advantage that all dorsal rim ommatidia can be stimulated simultaneously by a single stimulus positioned approximately in the optical center

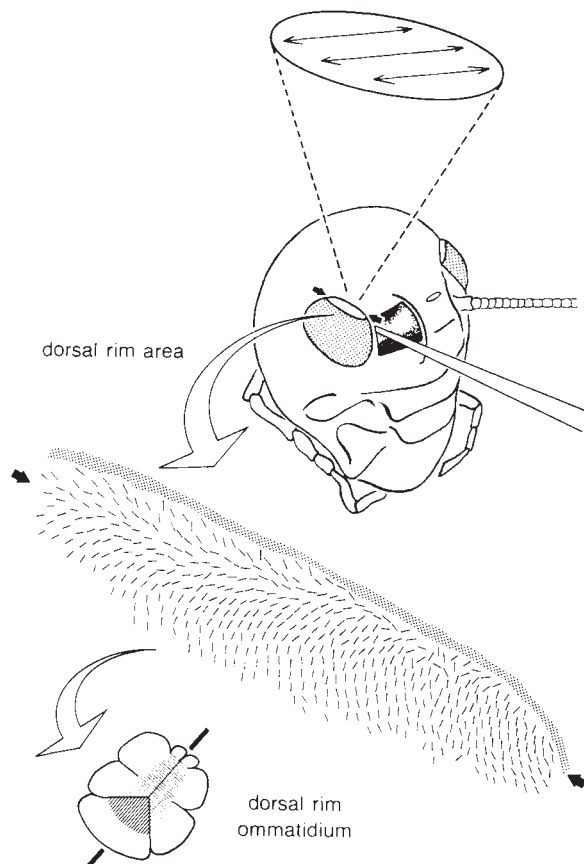


Fig. 1 Sketch of cricket head shows position of the specialized dorsal rim area of the eye (thick black arrows) and demonstrates method of stimulation and recording. In the expanded view of the dorsal rim area the orientations of transverse axes of the rhabdoms are indicated by bars. Note fan-shaped arrangement. Enlarged cross section of a dorsal rim ommatidium demonstrates the presence of two orthogonal microvillar directions within each rhabdom and defines the transverse axis of the rhabdom.

Methods. The dorsal rim area of field crickets, *Gryllus campestris*, was stimulated with polarized stimuli subtending visual angles of 1.5° or 50° and positioned 32° contralateral to the zenith (with respect to natural head position). For measuring the *e*-vector response (see Fig. 2 a-e) white light (Xenon arc) polarized with a HNPB polarizer (Polaroid) was used, giving 800 lux (1.5° stimulus) or 150 lux (50° stimulus) at the eye. Monochromatic light, as used for measuring spectral response and some of the intensity response functions, was produced with small-band interference filters (Schott). Intensity was adjusted with a quartz neutral-density wedge (Dyn Optics). Intra- and extracellular recordings from POL-neurons were obtained with KCl-filled glass micropipettes inserted into the medulla using standard electronic equipment and stored with an FM tape recorder. Response functions were determined off-line by counting the number of spikes during the last 800 ms of the 1,000-ms flashes with a window discriminator. *e*-vector response functions could also be obtained from experiments with continuously rotating *e*-vectors by counting spike numbers in 20° intervals. The $\Phi_{\max}$ and $\Phi_{\min}$ values were calculated by fitting second-order polynomials (parabolas) to the upper and lower arcs of the *e*-vector response functions. All *e*-vectors indicated in the paper refer to the long axis of the dorsal rim area of the right eye as specified in Fig. 4 (bottom).

of the dorsal rim area. Most neurons recorded responded strongly to changes of light intensity¹³ but were polarization-insensitive. This was in marked contrast to a number of neurons whose spike activity could be controlled by the orientation of the polarizer but were indifferent to intensity variations. We shall term these units POL-neurons.

POL-neurons exhibit spontaneous spike activity in the dark

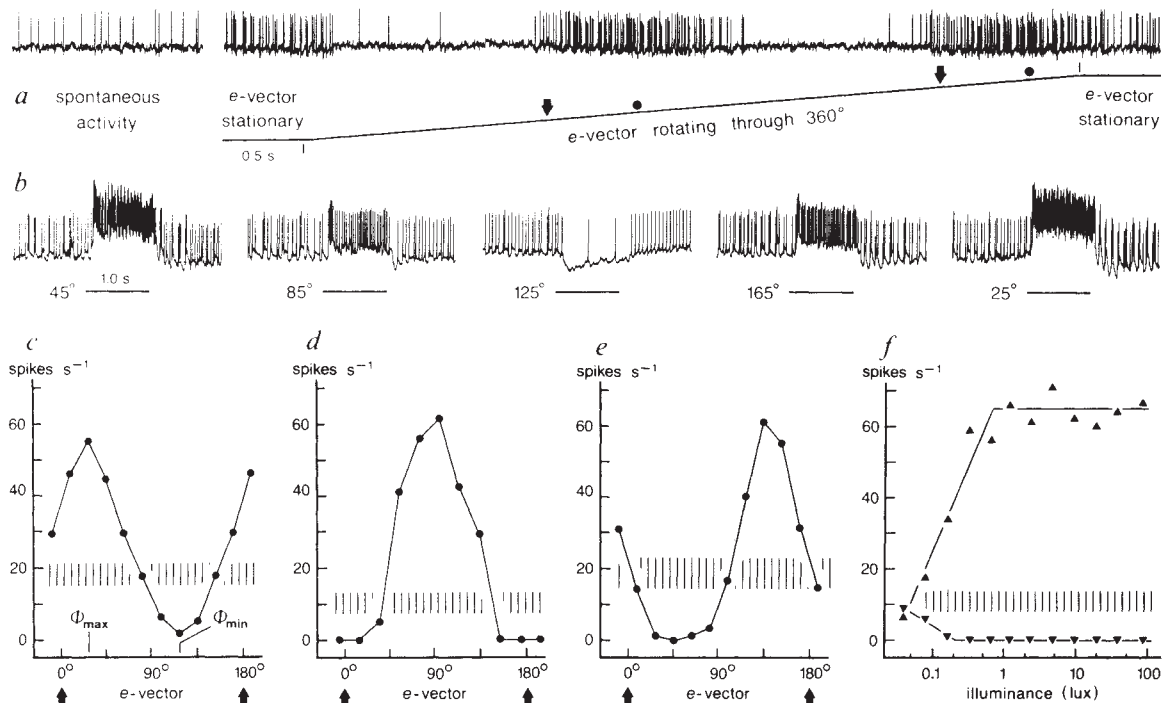


Fig. 2 Responses of six typical polarization-opponent neurons (POL-neurons): Note excitation and inhibition depending on *e*-vector orientation. *a*, Response to a continuously rotating *e*-vector. Arrows, *e*-vector parallel to long axis of dorsal rim area (compare Fig. 1 and Fig. 4 bottom); dots, $\Phi_{\max}$, that is, *e*-vector eliciting maximal spike frequency. Left trace shows spontaneous spike activity in the dark. *b*, Responses to 1-s flashes with different *e*-vectors, as indicated with respect to the long axis of the dorsal rim area; *c-e*, *e*-vector response functions as measured with 1-s flashes. Arrows, as in *a*. Hatched areas indicate range of spontaneous activity during the experiments in the dark (average $\pm$ s.d.). *f*, Intensity response functions with *e*-vectors near $\Phi_{\max}$ and $\Phi_{\min}$, respectively, using 1-s flashes. Note intensity independence of response above a critical light level. Stimulus conditions: white light (*a-f*), stimulus diameter 1.5° (*a, c, e*) or 50° (*b, d, f*). Calibrations: spike amplitude 5 mV (*a*) and 14 mV (*b*).

(range ~ 5 –25 spikes s^{-1}). With polarized light, spike frequency is a sinusoidal function of *e*-vector orientation with an excitatory and an inhibitory part. The response is strongly tonic. The *e*-vector eliciting maximal spike frequency ($\Phi_{\max}$; range 30–120 spikes s^{-1}) is oriented close to 90° ($88.8 \pm 2.1^\circ$, $n = 20$) to the *e*-vector of minimal spike frequency ($\Phi_{\min}$; often 0 spikes s^{-1}) (see Fig. 2*a-e*). Thus, POL-neurons are polarization-opponent neurons receiving antagonistic input from two analyser channels with orthogonal orientations of maximal sensitivity (Fig. 3). We propose that these two channels are represented by the two sets of receptors with orthogonally arranged microvilli present within each dorsal rim ommatidium (Fig. 1).

The *e*-vector response is independent of light intensity down to a rather low, critical light level (Fig. 2*f*). This is in accord with the model circuit given in Fig. 3 as the absolute response level of the analysers is irrelevant for the difference function. At lower light levels (corresponding to intensities normal for moonlight) the response drops steeply to zero within ~ 1 log unit of intensity ($n = 5$). Presumably this is because the receptor signals become unreliable due to quantum noise. With behavioural experiments, in which the cricket's spontaneous *e*-vector response was measured, intensity independence and a sharp drop of the response at low light levels was also observed (D. Herzmann & T.L., manuscript in preparation).

POL-neurons show very weak or no responses to flashes of unpolarized light although the two analyser channels within each dorsal rim ommatidium have different absolute sensitivities, as indicated by the different sizes ($\sim 2:1$) of the two microvillar blocks^{8,14} (see Fig. 1). Thus, their balanced inputs to the POL-neurons probably result from adjusting the synaptic gain.

Spectral response functions as measured in a few POL-neurons ($n = 3$) with monochromatic stimuli of sub-critical intensity exhibit a maximum in the blue range. This is in accord

with the findings that the dorsal rim ommatidia contain blue-receptors only (whereas other eye regions are composed of ultraviolet and green receptors)¹¹ and that the cricket's spontaneous *e*-vector response is most sensitive in the blue (D. Herzmann & T.L., manuscript in preparation).

Three types of POL-neuron characterized by different orientations of $\Phi_{\max}$ were found: their average $\Phi_{\max}$ values are $\sim 35^\circ$,

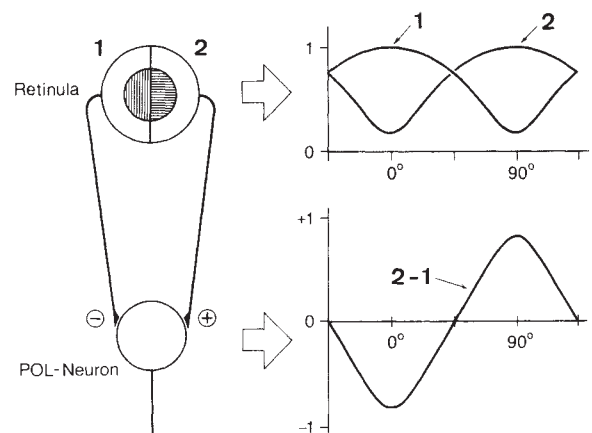


Fig. 3 Modelling POL-neurons. Left, schematic representation of a polarization opponent network. The two analyser channels (numbered 1 and 2), represented by the two receptors of a simplified dorsal rim retinula (compare Fig. 1), act antagonistically on a POL-neuron. Right, *e*-vector response functions of receptors (upper graph) and POL-neurons (lower graph). Ordinate, relative response; abscissa, *e*-vector orientation. The spike frequency of the POL-neuron follows the difference function (curve 2-1) as long as the cell potential does not fall below firing threshold.

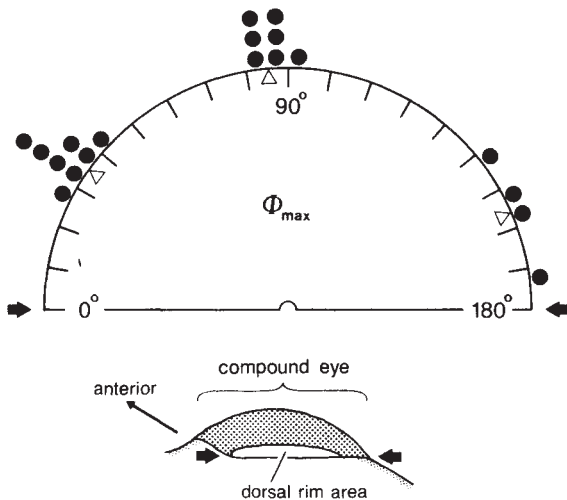


Fig. 4 Orientations of $\Phi_{\max}$ (e -vector eliciting maximal spike frequency) of different POL-neurons with reference to the long axis of the dorsal rim area. Dots, $\Phi_{\max}$ values of individual POL-neurons ($n = 19$); open triangles, average $\Phi_{\max}$ of each of the three classes of POL-neuron; arrows, long axis of dorsal rim area. Sketch (below) of a right eye as seen from the position of the stimulus (32° contralateral to the zenith) specifies reference direction.

85° and 155° with respect to the long axis of the dorsal rim area ($n = 19$) (Fig. 4). The orientation of the rhabdoms is not constant within the dorsal rim area but varies in a fan-like fashion (Fig. 1), suggesting that each type of POL-neuron receives input from a large number of ommatidia in three different parts of the dorsal rim area. Thus, e -vector information is not represented in fine spatial detail (as already indicated by the diffusing optics) but by just three types of interneuron, or even three individual neurons. What is their significance for the mechanism of e -vector detection? Two different models of e -vector detection, the logical framework of which was laid out earlier¹⁵, are presently discussed: in the 'simultaneous' model e -vectors are determined instantaneously by comparing the responses of three (or more) polarization-sensitive receptors with different analyser directions¹⁶. Thus, one could regard the three POL-neurons as 'macroreceptors' giving input into a three-dimensional simultaneous system. In the 'scanning' model an insect determines the symmetry plane of the celestial polarization pattern by scanning the sky with its dorsal rim areas, experiencing maximal response whenever the retinal array of receptors matches the e -vector pattern². Although one POL-neuron integrating e -vector information from the whole dorsal rim area would be sufficient to achieve this task, additional units directed forward and backward could serve to discriminate between solar and antisolar direction based on the different degrees of sky polarization.

The experiments of this study were not designed to provide conclusive evidence for one or the other model of e -vector detection. But we demonstrate that polarization-opponent mechanisms do occur in the insect visual system. These mechanisms could benefit the insect in two different respects: (1) making the system insensitive to intensity modulations and thus preventing interference of celestial intensity gradients with the e -vector signals; (2) increasing the amplitude of signal modulation, enhancing e -vector contrast.

I thank Drs S. Rossel and R. Wehner for enlightening discussions and critiques on the manuscript. Supported by the Swiss NSF.

Received 17 September; accepted 16 December 1987.

1. Rossel, S. & Wehner, R. *J. comp. Physiol. A* **154**, 607–615 (1984).
2. Rossel, S. & Wehner, R. *Nature* **323**, 128–131 (1986).
3. Leggett, L. M. W. *Nature* **262**, 709–711 (1976).

4. Waterman, T. H. & Aoki, K. *J. comp. Physiol.* **95**, 13–27 (1974).
5. Waterman, T. H. & Hashimoto, H. *J. comp. Physiol.* **95**, 1–12 (1974).
6. Wehner, R. & Strasser, S. *Physiol. Ent.* **10**, 337–349 (1985).
7. Brunner, D. & Labhart, T. *Physiol. Ent.* **12**, 1–10 (1987).
8. Burghause, F. M. H. R. *Zool. Jb. Physiol.* **83**, 502–525 (1979).
9. Sommer, E. W. thesis, Univ. Zürich (1979).
10. Labhart, T. *J. comp. Physiol.* **141**, 19–30 (1980).
11. Labhart, T., Hodel, B. & Valenzuela, I. *J. comp. Physiol. A* **155**, 289–296 (1984).
12. Meyer, E. P. & Labhart, T. *Cell Tissue Res.* **216**, 491–501 (1981).
13. Honegger, H.-W. *J. comp. Physiol.* **125**, 259–266 (1978).
14. Nilsson, D.-E., Labhart, T. & Meyer, E. P. *J. comp. Physiol. A* **161**, 645–658 (1987).
15. Kirschfeld, K. *Z. Naturforsch.* **27b**, 578–579 (1972).
16. Edrich, W. & Helversen v., O. *Biol. Cybern.* **56**, 89–96 (1987).

Activation of mitochondrial oxidative metabolism by calcium ions in *Limulus* ventral photoreceptor

Alan Fein*[†] & Marcos Tsacopoulos[‡]

* Laboratory of Sensory Physiology, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA

[‡] Experimental Ophthalmology Laboratory, 22, rue Alcide-Jentzer, 1211 Geneva 4, Switzerland

Cells regulate their metabolic energy production to meet the requirements of their energy consuming activities. For most animal cells the prime site of energy production, in the form of ATP, is the mitochondrion. Extensive *in vitro* studies of isolated mitochondria have provided detailed information about the specific biochemical reactions involved in energy production¹. At present there is a debate about whether respiration in excitable cells is controlled by the availability of ADP to the mitochondrion² and/or by calcium ions^{3–5}. Using the large ventral photoreceptor of the horseshoe crab (*Limulus polyphemus*) we describe a method for measuring the transient increase in the mitochondrial O_2 consumption (ΔQO_2) following a flash of light of a single photoreceptor. We then show that this ΔQO_2 results in part from a rise in the intracellular concentration of calcium (Ca_i).

Figure 1 shows the recording of a light-induced transient decrease in PO_2 (ΔPO_2), measured using a miniature O_2 electrode, in the extracellular region immediately surrounding the ventral photoreceptor cell body. This is the region where such a signal would be expected as it contains both visual pigment and mitochondria^{6,7}.

The experiment illustrated in Fig. 2 shows that the light-induced ΔPO_2 results from an increase in mitochondrial respiration⁸ and is not an artefact. An intracellular pipette was used to record the transmembrane potential of the photoreceptor and to inject rotenone, a specific inhibitor of oxidative metabolism⁹, while simultaneously recording the PO_2 using an O_2 electrode, as in Fig. 1. ΔQO_2 was calculated from ΔPO_2 as described in the legend to Fig. 2. It can be seen in Fig. 2e that rotenone strongly inhibited both phases of the ΔQO_2 without affecting the sensitivity of the cell, or inhibiting the light-induced receptor potential (Fig. 2a and c). We conclude that the light-induced ΔPO_2 results from activation of mitochondrial respiration.

The light-induced rise in Ca_i (ΔCa_i) was measured (see Fig. 3a) with a double-barrel calcium ion-selective electrode while simultaneously measuring the local PO_2 , as in Figs 1 and 2. The similarity between the time courses of ΔCa_i and ΔQO_2 is quite striking and suggests that light produces a ΔCa_i that is both large and rapid enough to activate oxidative metabolism. If this activation of oxidative metabolism is the result of the light-induced ΔCa_i , then calcium chelators injected into the cell should suppress the light-induced ΔQO_2 . Such an effect is illustrated in Fig. 3b. In this experiment an intracellular pipette was used to inject the calcium chelator EGTA into the cell and the

[†] Present address: Department of Physiology, University of Connecticut Health Center, Farmington, Connecticut 06032-9984, USA.

Fig. 1 Light-induced oxygen consumption of a single ventral photoreceptor. The cell in *a*, *b* and *c* is being held away from the ventral nerve by suction applied to the large blunt pipette shown at the left. This suction pipette was used to strip the cell of its glia and connective tissue¹⁷. At the right is the miniature platinum O₂ electrode used to measure the PO₂ in the extracellular space surrounding the cell. In *a*, *b* and *c*, the O₂ electrode was placed at three different locations along the cell, and in each instance the cell was stimulated by the same intensity and duration light flash. The recordings were obtained in the order shown. After obtaining the recording in *c*, the O₂ electrode was returned to the position in *a*, and the same stimulus as in *a*, *b* and *c* elicited a response equal in amplitude to that in *a*. The stimulus flash used in this and the other experiments reported here had a duration of 1 s and was attenuated 100-fold from the maximum intensity available from the xenon arc light source.

Methods. The preparation is the same as the one first described by Millecchia and Mauro¹⁸. The excised ventral nerve was pinned down in a large chamber and covered with a 1–2-mm-thick layer of artificial sea water. The photoreceptor was photographed on the video monitor on which the image was projected by means of an infrared video camera¹⁹. The tip of the miniature O₂-sensitive electrode can be seen to the right of the cell. The electrode has a recess between the open tip and the measuring surface of the platinum wire, visible in *a*. Because of its design and small size this type of electrode can measure localized rapid changes in PO₂ (refs 8, 20). We used electrodes that were chosen for their stability and low noise; their O₂ current ranged from 0.35 to 0.55 pA per mm Hg of PO₂.

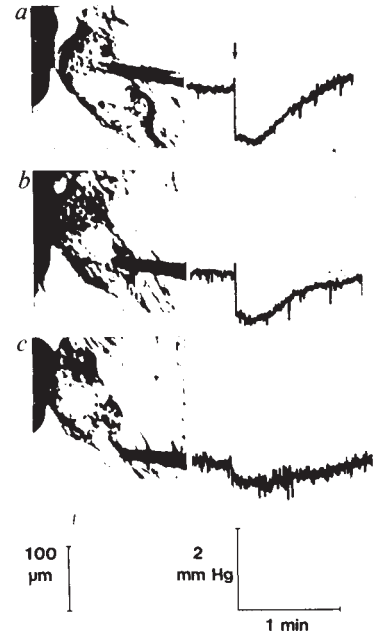
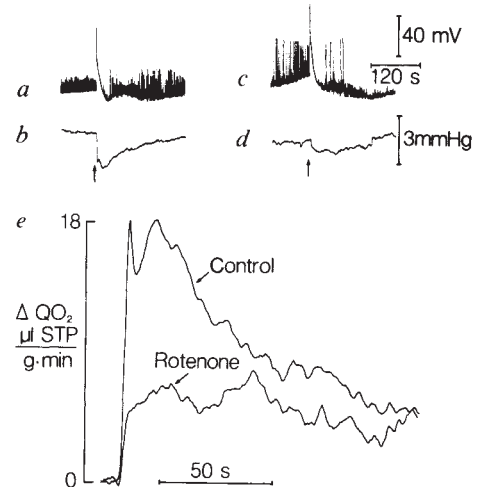


Fig. 2 Inhibition of oxidative metabolism by rotenone. *a*, *c*, Transmembrane potential measured by an intracellular pipette containing a carrier solution of 100 mM potassium aspartate, 10 mM HEPES, pH 7.0, to which rotenone dissolved in ethanol was added to give a final concentration of 10 μ M rotenone and 1% ethanol. *b*, *d*, PO₂ recordings obtained using a miniature O₂ electrode as in Fig. 1. *e*, Δ QO₂ calculated from the Δ PO₂ records in *b* and *d* as indicated. Records *a* and *b* were obtained before, and *c* and *d* after, injection of rotenone. Injection of the 1% ethanol carrier solution without rotenone did not inhibit the Δ QO₂. The arrows indicate the occurrence of the light flash which elicited the receptor potential and corresponding Δ PO₂.

Methods. Δ QO₂ was calculated from the Δ PO₂ recording by considering the photoreceptor as a sphere²¹ of 50 μ m radius that consumes O₂ uniformly throughout. The diffusion (*D*) and solubility (α) coefficients of O₂ in the cell were assumed to be the same as in the bath ($D = 2 \times 10^{-5}$ cm² s⁻¹, $\alpha = 31$ μ l O₂ STP (standard temperature and pressure) cm⁻³ atm⁻¹); tests of the method have shown that plausible errors in *D*, α and radius can affect the amplitude of Δ QO₂, but the kinetics of Δ QO₂ only negligibly (S. Poitry and H. Widmer, personal communication). Samples of the PO₂ record were taken every 400 ms. The Fourier transform of the Δ PO₂ was computed using a fast Fourier transform (FFT) routine; the result was multiplied by the transfer function obtained by solving the equation of diffusion for the case of a sphere consuming a substance uniformly, and the inverse Fourier transform of the product then computed with the FFT routine to yield the Δ QO₂. For convenience, the values are connected with a continuous line. Both the Δ PO₂ and Δ QO₂ before injection of rotenone are biphasic, consisting of an early rapid phase followed by a later slower phase. There is little difference between the time courses of the Δ QO₂ and the resulting Δ PO₂, probably because any region of the cell that consumes O₂ is less than 50 μ m away from the bath which supplies O₂, so that diffusion of O₂ does not introduce as much of a delay between Δ QO₂ and Δ PO₂ as in other preparations^{8,20}. The transmembrane potential recordings (*a* and *c*) exhibit discrete membrane depolarizations, some of which occur spontaneously and others in response to the absorption of stray light, these discrete depolarizations have been termed 'quantum bumps' as they seem to result from the activation of the cell by single photons²². The amplitude of the quantum bumps is a good indication of the sensitivity of the cell²³. Rotenone injection does not significantly reduce the amplitude of the bumps. We integrated the area under the Δ QO₂ to obtain the total amount of oxygen consumed; rotenone inhibited the O₂ consumption by 68%. In three other cells, rotenone inhibited the light-induced Δ PO₂ by more than 80%. The amount of rotenone required for complete inhibition of respiration *in vitro* is 30 nmol g⁻¹ mitochondrial protein⁹.



PO₂ was measured as in Figs 1 and 2. Figure 3*b* shows the superimposed Δ QO₂ obtained before and after injection of EGTA. It can be seen that the EGTA completely blocks the rapid phase of Δ QO₂ and inhibits the slower phase, indicating that an increase in Ca_i is necessary for the rapid phase and contributes towards the slower phase.

If calcium has a function in activating mitochondrial respiration, as the results in Fig. 3 indicate, then oxygen consumption should be activated by directly injecting calcium into the cell.

In the experiment illustrated in Fig. 4, a cell was impaled with a pipette containing 20 mM CaCl₂ and the PO₂ was measured as in Figs 1–3. It can be seen that the first light flash elicited a typical receptor potential (Fig. 4*a*) and Δ PO₂ (Fig. 4*b*). Subsequently, injection of calcium resulted in a transient depolarization (Fig. 4*a*) and a Δ PO₂ consisting of two components: a large rapid component, followed by a smaller, slow component, as after a light flash (Fig. 4*b*). The cell was not damaged by the injection because another flash after the calcium injection again

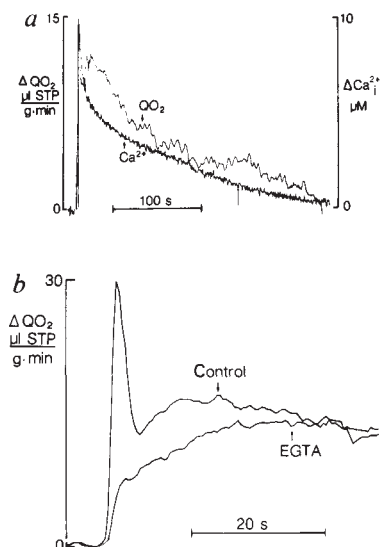


Fig. 3 *a*, Simultaneous measurement of ΔCa_i and ΔQO_2 in a single ventral photoreceptor. The cell was impaled with a double-barrel quartz calcium ion selective electrode²⁴. Ca_i was recorded as the voltage difference between the two barrels of the calcium-ion selective electrode, calibrated as described²⁵. The ΔQO_2 was computed from the PO_2 recording as described in Fig. 2.

Methods. Illumination of ventral photoreceptors results in an elevation of Ca_i (ref. 26) that is not uniformly distributed throughout the photoreceptor cell body²⁷. The cell body of a ventral photoreceptor is typically segmented into two lobes, the light-sensitive rhabdomeral (R) lobe, where phototransduction occurs, and an insensitive arhabdomeral (A) lobe, which contains the nucleus^{7,17}. The R-lobe typically consists of the portion of the cell body most distal from the axon. The light-induced ΔCa_i is initiated in the R-lobe, consequently, to detect a large rapid rise in Ca_i , the calcium ion-selective electrode must be placed in the R-lobe²⁵. If this electrode is placed in the A-lobe, a small and delayed rise in Ca_i is detected, consistent with diffusion of calcium from the R- to A-lobe²⁵. The precise distribution of mitochondria between the two lobes is unknown, so at present we cannot propose a quantitative relationship between ΔCa_i and ΔQO_2 , hence we compare the time course of ΔCa_i with the calculated ΔQO_2 with both arbitrarily scaled to equal peak amplitude. Because of the slow response time of the microelectrode, the recorded Ca_i appears to rise more slowly than it actually does. From these ion-selective electrode recordings we cannot say whether the ΔCa_i precedes the ΔQO_2 , but we have found by using the calcium-sensitive luminescent protein aequorin that the Ca_i rises well before the rise in QO_2 after a flash of light (results not shown). *b*, Inhibition of light-induced ΔQO_2 by the calcium chelator EGTA. The cell was impaled by a pipette containing 100 mM K₂ EGTA, 95 mM Ca(OH)₂ neutralized to pH 7.0.

Methods. Comparison of the ΔQO_2 shown here with those in Figs 2 and 4 correctly shows that there is a significant variation from cell to cell in the amplitude of both the rapid and slow components of ΔQO_2 . For 20 cells, the peak amplitude of the rapid phase was $28.3 \pm 1.85 \mu\text{l STP per g min}^{-1}$ (mean $\pm$ s.d.) and the peak amplitude of the slow phase was 15.2 ± 5.52 . We pressure-injected a mixture of EGTA and calcium (Ca/EGTA ratio 0.95), rather than EGTA alone, to prevent any drastic decrease in resting level of Ca_i in the cells (pp. 826–828 of ref. 25). To establish that the effect resulted from the EGTA and not the calcium in the injection solution, the calcium concentration was reduced from 95 mM to 20 mM to give a Ca/EGTA ratio of 0.2. Injections of this solution produced similar effects. The example given is the smallest inhibition of the area under the ΔQO_2 (11%) observed, between 30 and 60% inhibition being observed in the other five cells so injected. Another 14 cells were injected with calcium buffers and the light-induced ΔPO_2 recorded, but the ΔQO_2 could not be calculated without an on-line computer. In eight cells injected with the EGTA solutions and six cells injected with 100 mM of the calcium chelator BAPTA (Ca/BAPTA ratio of 0.5), there was complete suppression of the rapid phase and reduction in the amplitude of the slow phase of the light-induced ΔPO_2 .

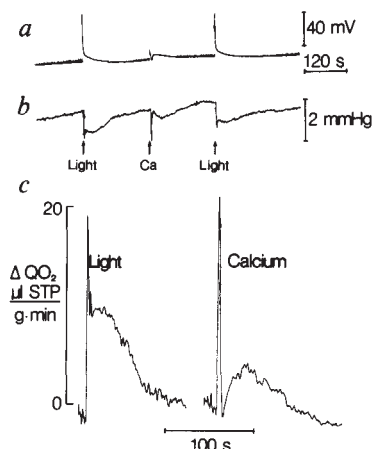


Fig. 4 Activation of oxidative metabolism by calcium injection. *a*, Transmembrane potential recorded by an intracellular pipette containing 100 mM potassium aspartate, 20 mM CaCl₂, 10 mM HEPES, 0.4% Triton X-100, pH 7.0. *b*, PO_2 recordings obtained using a miniature O_2 electrode as in Figs 1–3. *c*, ΔQO_2 calculated from the ΔPO_2 records in *b* for the first light flash and calcium injection. Results similar to these were obtained in four other cells. The arrows show the occurrence of the light flash and the bar indicates the injection of calcium into the cell. The injection solution contained 0.4% Triton X-100 because Triton prevented clogging of the calcium-injection electrode²⁸. Control injections (not shown) of the above solution without calcium did not activate O_2 consumption in six cells.

Methods. To mimic the rise in Ca_i induced by light, calcium must be injected into the R-lobe (ref. 28), but even if injected directly into the R-lobe, there are still substantial differences in the spatial distribution of the resulting rises in Ca_i . Illumination results in increased calcium throughout the R-lobe, which can typically have a diameter of 50 μm (ref. 29), whereas pressure-injection of 1–2 mM calcium into the R-lobe only raises calcium within a subregion of the R-lobe centred around the injection site²⁹. When we injected solutions containing 1–2 mM calcium into the R-lobe of ventral photoreceptors (ten cells) while recording the PO_2 , we failed to detect a consistent change in PO_2 . To more effectively elevate the calcium throughout the R-lobe, we pressure-injected solutions containing 20 mM calcium into the photoreceptors.

elicited a receptor potential and a ΔPO_2 . In Fig. 4c we compare the ΔQO_2 induced by light and calcium injection; the similarities in the two time-courses are striking.

In mammalian mitochondria, calcium could stimulate respiration by activation of intramitochondrial dehydrogenases¹⁰. To test whether calcium was acting at an intramitochondrial site, we injected ventral photoreceptors with ruthenium red, a potent inhibitor of calcium uptake into isolated mitochondria¹¹. Cells were injected to a final concentration in excess of $1 \mu\text{g ml}^{-1}$, which for this sample of ruthenium red blocks calcium uptake *in vitro* (R. Denton, personal communication). A reproducible inhibition of either phase of the light-induced ΔPO_2 following such injections (results not shown) was not observed. This finding is consistent with the report that mitochondria from non-vertebrate sources do not have intramitochondrial enzymes sensitive to calcium¹⁰.

Illumination of ventral photoreceptors causes an influx of sodium ion which raises the concentration of intracellular sodium ion and thereby activates the sodium pump¹². The pump uses energy released by splitting ATP to pump sodium out of the cell. It has been suggested that in neurons the associated increase in ADP acts as a stimulus for mitochondrial respiration¹³. We have found that pressure-injection of sodium into ventral photoreceptors in amounts sufficient to raise the concentration of intracellular sodium above that obtained with light does not activate O_2 consumption (results not shown). Thus, activity of the sodium pump does not appear to stimulate

mitochondrial respiration, confirming previous results in bee photoreceptors¹⁴.

Our findings strongly support the hypothesis that a rise in Ca_i activates mitochondrial respiration. In summary, we have shown that there is a light-induced ΔCa_i that has a similar time-course to the ΔQO_2 that EGTA tends to block the light-induced ΔQO_2 , and that pressure injection of calcium into the cell induces a ΔQO_2 .

Calcium has been shown to activate the oxidation of α -glycerolphosphate by insect flight muscle¹⁵, where the calcium-sensitive mitochondrial enzyme glycerolphosphate dehydrogenase is located at the external face of the inner mitochondrial membrane, making it accessible to cytosolic calcium¹⁶. We suggest that activation of the dehydrogenase by a light-induced rise in Ca_i results in an increased rate of mitochondrial respiration in *Limulus* ventral photoreceptors.

This work was supported by the Swiss NSF, the Georges Kern Foundation, the Zyma Foundation and the National Institutes of Health. We particularly thank S. Poitry for writing computer programs used in this study and for critical discussion. We also thank Drs R. Payne, C. Wolheim, W. Schlegel and Ms H. Widmer for helpful comments on the manuscript, J.-L. Munoz and P. Perrotet for making microelectrodes and D. Pollet for construction of apparatus.

Lithium inhibits adrenergic and cholinergic increases in GTP binding in rat cortex

Sofia Avissar*, Gabriel Schreiber†, Abraham Danon* & R. H. Belmaker†‡

* Clinical Pharmacology Unit. † Beer Sheva Mental Health Center, Ida and Solomon Stern Psychiatry Research Unit, Ben Gurion University of the Negev, P.O.B. 4600, Beer Sheva, Israel

Lithium is a unique drug with therapeutic as well as prophylactic value for both manic and depressive phases of manic-depressive illness^{1,2}. The precise mechanisms of its clinical efficacy remain unknown, but there are two main theories of its biochemical action. One proposes that lithium inhibits adrenergically activated adenylate cyclase function³⁻⁵ whereas the other suggests that it inhibits phosphatidyl inositol turnover, which is known to be activated by cholinergic agonists⁶⁻⁸. Neither mechanism alone, however, can explain both the antimanic and antidepressant effects of lithium. Because of the pivotal role of G proteins in post-receptor information transduction, we have investigated the interaction of lithium with G protein function. Lithium at therapeutically efficacious concentrations completely blocked both adrenergic and cholinergic agonist-induced increases in [³H]GTP binding to membranes from rat cerebral cortex, in both *in vitro* and *ex vivo* experiments. The same lithium treatments also abolished guanine nucleotide modulation of agonist binding. Our findings suggest G proteins (G_s and G_i or G_o) as the molecular site of action for both the antimanic and antidepressant effects of lithium.

An important characteristic of G proteins is their increased guanine nucleotide binding following agonist stimulation which in turn leads to their activation. Tritiated guanine nucleotide binding stimulated by agonist is an established test for G-protein function. It has been used in membrane preparations^{9,10}, and also for the purification of G proteins from crude membrane preparations¹¹ and has been both kinetically and mechanistically analysed using various purified G proteins¹² in reconstituted systems^{13,14}. Although basal binding of guanine nucleotides to certain crude membrane preparations is not exclusive to G

Received 21 October; accepted 1 December 1987.

- Nichols, D. G. *Bioenergetics: an Introduction to the Chemiosmotic Theory* (Academic, New York, 1982).
- Chance, B. & Williams, G. R. *Adv. Enzymol.* **17**, 65-134 (1956).
- Denton, R. M. & McCormack, J. G. *FEBS Lett.* **119**, 1-8 (1980).
- Hansford, R. G. *Rev. Physiol. Biochem. Pharmac.* **102**, 1-72 (1985).
- Denton, R. M. & McCormack, J. G. *Cell Calcium* **7**, 377-386 (1986).
- Clark, A. W., Millecchia, R. & Mauro, A. J. *gen. Physiol.* **54**, 289-309 (1969).
- Calman, B. G. & Chamberlain, S. C. *J. gen. Physiol.* **80**, 839-862 (1982).
- Tsacopoulos, M. & Poitry, S. *J. gen. Physiol.* **80**, 19-55 (1982).
- Slater, E. C. *Meth. Enzym.* **10**, 48-57 (1967).
- McCormack, J. G. & Denton, R. M. *Biochem. J.* **196**, 619-624 (1981).
- Carafoli, E. & Sacktor, B. *Biochem. biophys. Res. Commun.* **49**, 1489-1503 (1972).
- Brown, J. E. & Lisman, J. E. *J. gen. Physiol.* **59**, 720-733 (1973).
- Ritchie, J. M. *Prog. Biophys. molec. Biol.* **26**, 147-187 (1973).
- Tsacopoulos, M., Orkand, R. K., Coles, J. A., Levy, S. & Poitry, S. *Nature* **301**, 604-606 (1983).
- Hansford, R. G. & Chappel, J. B. *Biochem. biophys. Res. Commun.* **27**, 686-692 (1967).
- Klingenberg, M. & Buchholz, M. *Eur. J. Biochem.* **13**, 247-252 (1970).
- Stern, J., Chinn, J., Bacigalupo, J. & Lisman, J. E. *J. gen. Physiol.* **80**, 825-837 (1982).
- Millecchia, R. & Mauro, A. J. *gen. Physiol.* **54**, 310-330 (1969).
- Corson, D. W. & Fein, A. *Biophys. J.* **44**, 299-304 (1983).
- Tsacopoulos, M., Poitry, S. & Borsellino, A. *J. gen. Physiol.* **77**, 601-628 (1981).
- Poitry, S. & Widmer, H. *J. Physiol., Lond.* **378**, 65p (1986).
- Yeandle, S. & Spiegler, J. B. *J. gen. Physiol.* **61**, 552-571 (1983).
- Wong, F. *Nature* **276**, 76-78 (1978).
- Coles, J. A. & Orkand, R. K. *J. Physiol., Lond.* **362**, 415-435 (1985).
- Levy, S. & Fein, A. *J. gen. Physiol.* **85**, 805-841 (1985).
- Brown, J. E. & Blinks, J. R. *J. gen. Physiol.* **64**, 643-665 (1974).
- Harary, H. & Brown, J. E. *Science* **224**, 292-293 (1984).
- Payne, R., Corson, D. W. & Fein, A. *J. gen. Physiol.* **88**, 107-126 (1986).
- Payne, R. & Fein, A. *J. Cell Biol.* **104**, 933-937 (1987).

Table 1 Effects of cholera toxin- or pertussis toxin-catalysed ADP-ribosylation on agonist-induced increase in GTP binding

Agonist	Per cent increase in [³ H]GTP-binding capacity		
	Membrane treatment		
	None	Cholera toxin	Pertussis toxin
Isoprenaline (25 μM)	21.8 $\pm$ 2.2	0	21.1 $\pm$ 2.0
Carbachol (50 μM)	21.6 $\pm$ 3.2	21.1 $\pm$ 2.5	8.6 $\pm$ 1.0

Cortical membranes (8.6 mg protein) were suspended in 1 ml buffer containing 25 mM Tris pH 7.4, 10 mM NAD, 1 mM ATP, 1 mM EGTA, 100 μM GTP, 2 mM dithiothreitol (DTT), 5 mM MgCl_2 , 10 mM thymidine, 20 mM creatine phosphate and 40 $\mu\text{g ml}^{-1}$ creatine phosphokinase. ADP-ribosylation was carried out for 15 min at 30 °C by adding cholera toxin (50 $\mu\text{g ml}^{-1}$), pre-activated for 10 min at 37 °C with 20 mM DTT or pertussis toxin (25 $\mu\text{g ml}^{-1}$) pre-activated for 10 min at 30 °C with 20 mM DTT. The reaction was stopped by the addition of 25 ml ice-cold 25 mM Tris pH 7.4 and 5 mM MgCl_2 immediately followed by centrifugation at 10,000g for 10 min. [³H]GTP binding was then carried out as described in the legend to Fig. 1.

proteins⁹, there is no dispute that the agonist-induced increase in tritiated guanine nucleotide binding is an exclusive characteristic of G proteins¹⁵.

The effect of the β -adrenergic agonist, isoprenaline, on [³H]GTP-specific binding to membranes of rat cerebral cortex is shown in Fig. 1a. The isoprenaline-induced increase in GTP binding was abolished by the β -adrenergic antagonist, propranolol (Fig. 1a). As the β -adrenergic system has been shown to increase intracellular cyclic AMP (cAMP) levels through activation of G_s ^{16,17}, our findings reflect an increase in G_s -binding capacity. To substantiate this conclusion, ribosylation of G_s catalysed by cholera toxin was carried out before [³H]GTP binding experiments. A complete abolition of the isoprenaline-induced increase in GTP binding was obtained following ribosylation of G_s (Table 1). Isoprenaline increases GTP binding selectively by acting through G_s , as ribosylation of other G proteins catalysed by pertussis toxin did not affect its action (Table 1). Lithium at therapeutically effective concentrations (0.6 mM) completely blocked isoprenaline effects on GTP binding to G_s (Fig. 1a), with no effect on basal GTP binding.

Similarly, increases in GTP binding capacity induced by the muscarinic agonist, carbachol, and blockable by the muscarinic

‡ To whom correspondence should be addressed.

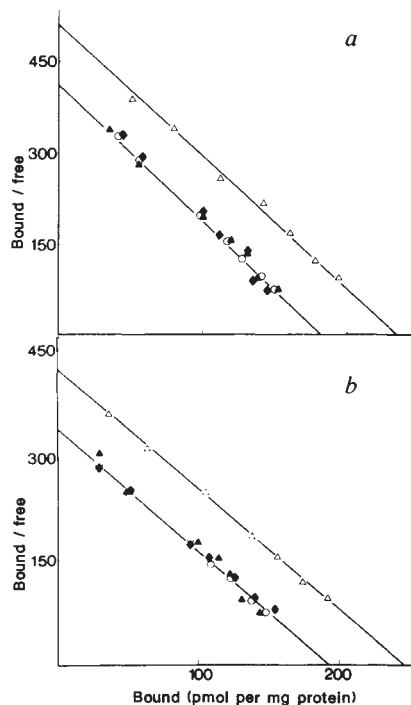


Fig. 1 $[^3\text{H}]\text{GTP}$ binding to rat cerebral cortical membranes. *a*, Effect of isoprenaline on $[^3\text{H}]\text{GTP}$ binding. *b*, Effect of carbachol on $[^3\text{H}]\text{GTP}$ binding. 200 μl of cortical membranes (0.5–1.5 mg protein) from Charles-River male rats suspended in 25 mM Tris (pH 7.4), 1 mM ATP, 0.1 mM Mg^{2+} (same results were obtained using various concentrations of Mg^{2+} reaching 5 mM), 1 mM EGTA and 1 mM DTT, were pipetted into plastic microfuge tubes containing varying concentration of $[^3\text{H}]\text{GTP}$ (0.05–5 μM). Incubation was carried out at room temperature for 15 min (equilibrium conditions) and the reaction was stopped by the addition of 5 volumes of ice-cold buffer A (25 mM Tris pH 7.4, 1 mM DTT), followed by centrifugation at 10,000g for 1 min. The pellets were rapidly washed three times with ice-cold buffer A by repeating centrifugation and then resuspended in 200 μl of the same buffer. Bound radioactivity was measured by liquid scintillation spectrometry by the addition of a 150 μl aliquot dissolved in scintillation liquid. All assays were carried out in triplicate, together with triplicate control samples containing 100 μM unlabelled Gpp(NH)p for determination of non-specific binding. *a*, Isoprenaline was added at a final concentration of 25 μM , DL-propranolol 10 μM and lithium 0.6 mM. The concentration of isoprenaline was the minimal which induced maximal agonist effect (dose-response curve not shown). (O) basal specific GTP binding; (Δ) in the presence of isoprenaline; ($\blacktriangle$) in the presence of isoprenaline and propranolol; ($\blacklozenge$) in the presence of isoprenaline and lithium. Basal maximal binding capacity was 185 ± 11.7 pmoles per mg protein. Maximal binding capacity in the presence of isoprenaline was 231.1 ± 12.4 pmoles per mg protein. *b*, (O) basal $[^3\text{H}]\text{GTP}$ -specific binding; (Δ) in the presence of 50 μM carbachol; ($\blacktriangle$) in the presence of 50 μM carbachol and 5 μM atropine; ($\blacklozenge$) in the presence of 50 μM carbachol and 0.6 mM lithium. The concentration of carbachol was the minimal which induced the maximal agonist effect (dose-response curve not shown). Maximal binding capacity in the presence of carbachol was 249.6 ± 15.2 pmoles per mg protein. Addition of 100 mM NaCl instead of lithium did not abolish the agonist effects. $P < 0.01$ for $[^3\text{H}]\text{GTP}$ maximal binding capacity in the presence of isoprenaline or carbachol as compared with control.

antagonist, atropine, were found to be abolished by lithium (Fig. 1*b*). Because the muscarinic system is known to be coupled to G proteins other than G_s , that is G_i and G_o ^{14,18}, lithium abolishment of the carbachol effect on GTP binding reflects its interaction with the latter G proteins. Ribosylation of G proteins other than G_s , catalysed by pertussis toxin, inhibited the carbachol-induced increase in GTP binding capacity (Table 1). But ribosylation catalysed by cholera toxin did not affect carbachol action

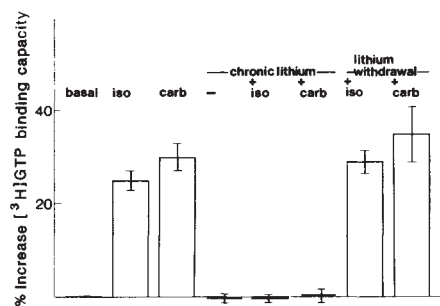


Fig. 2 Effects of isoprenaline and carbachol on $[^3\text{H}]\text{GTP}$ binding in chronically lithium-treated rats. Charles-River male rats (2–3 months old, 200–250 g) were i.p. injected with 25 mg kg^{-1} Li_2CO_3 daily for 12–21 days. Control rats were injected daily with saline. Membrane preparation as well as binding-experimental procedures were the same as described in the legend to Fig. 1. Isoprenaline concentration was 25 μM , carbachol was 50 μM .

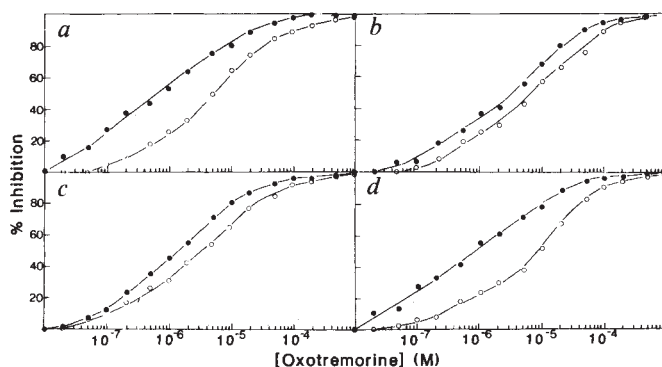


Fig. 3 Lithium inhibition of Gpp(NH)p effect on agonist binding to muscarinic receptors in rat brain medulla. Inhibition-concentration curves for the muscarinic agonist oxotremorine, binding in the absence ($\bullet$) or presence ($\circ$) of 75 μM Gpp(NH)p. *a*, Modulation of oxotremorine binding by Gpp(NH)p. *b*, *In vitro* effect of lithium (LiCl 1.5 mM) on Gpp(NH)p modulation of oxotremorine binding. *c*, *Ex vivo* effect of lithium on Gpp(NH)p modulation of oxotremorine binding (conditions of chronic lithium treatment are described in the legend to Fig. 2). *d*, Abolishment of the *in vitro* lithium effect on Gpp(NH)p modulation of oxotremorine binding by ribosylation of G proteins catalysed by pertussis toxin (see legend to Table 1).

Methods. Oxotremorine binding to membranes of rat brain medulla was determined through competition with $[^3\text{H}]\text{quinuclidinyl benzilate}$ ($[^3\text{H}]\text{QNB}$). The competition experiments were conducted using $[^3\text{H}]\text{QNB}$ at a fixed concentration of 0.5 nM in the absence or presence of various concentrations of oxotremorine. The reaction was started by adding the membranes (0.3 mg ml^{-1}) to the reaction mixture which contained the labelled ligand and the unlabelled agonist in 0.5 ml Krebs-Henseleit buffer pH 7.4 (25 mM Tris Cl, 118 mM NaCl, 11.1 mM glucose, 4.7 mM KCl, 1.9 mM CaCl_2 , 1.0 mM NaH_2PO_4 , 1.1 mM MgCl_2). The incubation was carried out for 45 min at 22 $^{\circ}\text{C}$. The reaction was stopped by the addition of 3 ml ice-cold buffer and filtering through GF/C filters (Whatman). The filters were subsequently washed twice with 3 ml of cold buffer and taken for scintillation counting. Non-specific binding (in the presence of 10 μM atropine) was subtracted in all cases.

(Table 1), thus supporting a muscarinic effect through G proteins other than G_s .

The *in vitro* effects of lithium on G proteins were substantiated by *ex vivo* experiments in chronically lithium-treated male rats. Chronic treatment with daily intraperitoneal (i.p.) injections of lithium carbonate for 12–21 days, reaching therapeutic blood levels (0.6–1.0 mM) prevented both isoprenaline and carbachol-induced increases in GTP binding capacity (Fig. 2). No differ-

ences in basal GTP binding characteristics were found between lithium-treated and control saline-treated rats (Fig. 2). Withdrawal of lithium for 48 h resulted in full recovery of both isoprenaline- and carbachol-induced increases in GTP binding (Fig. 2).

The modulation by guanine nucleotides of agonist binding to muscarinic receptors is well described¹⁹⁻²¹ and has been proven to be mediated through either G_i or G_o ¹⁸. We have used Gpp(NH)p-dependent changes in muscarinic agonist-binding characteristics (Fig. 3a) to muscarinic receptors in membranes from rat brain medulla as an alternative method for studying the effect of lithium on G-protein function. Both *in vitro* and *ex vivo* treatments with lithium at therapeutically effective concentrations inhibited the ability of Gpp(NH)p to modulate muscarinic agonist binding (Figs 3b and c). Ribosylation catalysed by pertussis toxin totally abolished lithium *in vitro* effects on Gpp(NH)p modulation of the muscarinic agonist binding (Fig. 3d).

Various studies have shown attenuation by lithium of either β -agonist induced, Gpp(NH)p induced or forskolin-induced activation of adenylate cyclase²². The β -agonist and the guanine nucleotide clearly act through G proteins. Recent work has revealed two binding sites for forskolin in rat brain. A low-affinity binding site is located on the catalytic unit, whereas a high-affinity binding site is thought to be associated with the complex of the catalytic unit and the activated G protein²³. Thus lithium attenuation of forskolin-induced adenylate cyclase activation may also be mediated through an effect on G proteins. Recently, it was found that lithium produced an inhibition, rather than enhancement, of the inositol trisphosphate ($InsP_3$) response to both muscarinic and α -1-adrenergic agonists in the rat cortex^{24,25}. This inhibition could not be explained by the classical, inositol-1-phosphatase, site of lithium effects on phosphatidyl inositol (PtdIns) metabolism. The accumulating data on the regulatory role played by G proteins in PtdIns metabolism²⁶⁻²⁸ suggests that this additional lithium site might also be located on G proteins. Since Janowsky and colleagues²⁹ formulated an adrenergic-cholinergic balance hypothesis of bipolar disorders, substantial evidence has accumulated supporting this hypothesis. Our present findings on the effects of lithium on both G_s and G_i or G_o function suggest G proteins as the common site for both the antimanic and the antidepressant therapeutic effects of lithium, which may act by competing with Mg^{2+} ions, known to be essential for GTP binding to G proteins.

S.A. is a Rachela Prywes Fellow of the Israel Institute for Psychobiology Charles E. Smith Family Foundation.

Received 24 September; accepted 7 December 1987.

1. Schou, M. & Thomsen, K. in *Lithium Research and Therapy* (ed. Johnson, F. N.) 63-84 (Academic, New York, 1975).
2. Schildkraut, J. in *Lithium, Its Role in Psychiatric Research and Treatment* (eds Gershon, S. & Shopsin, B.) 51-53 (Plenum, New York, 1973).
3. Forn, J. in *Lithium Research and Therapy* (ed. Johnson, F. N.) 487-497 (Academic, New York, 1975).
4. Ebstein, R., Belmaker, R. H., Grunhaus, L. & Rimon, R. *Nature* **259**, 411-413 (1976).
5. Belmaker, R. H. *Biol. Psychiat.* **16**, 333-349 (1981).
6. Hallcher, L. M. & Sherman, W. R. *J. biol. Chem.* **255**, 10896-10901 (1980).
7. Berridge, M. J., Downes, C. P. & Hanley, M. R. *Biochem. J.* **206**, 587-595 (1982).
8. Allison, J. H., Blisner, M. E., Holland, W. H., Hipps, P. P. & Sherman, W. R. *Biochem. biophys. Res. Commun.* **71**, 664-669 (1987).
9. Cassel, D. & Selinger, Z. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4155-4159 (1978).
10. Murayama, T. E. & Ui, M. *J. biol. Chem.* **259**, 761-769 (1984).
11. Sternweis, P. C. & Robishaw, J. D. *J. biol. Chem.* **259**, 13806-13813 (1984).
12. Fergusson, K. M., Higashijima, T., Smigel, M. D. & Gilman, A. G. *J. biol. Chem.* **261**, 7393-7399 (1986).
13. Northup, J. K., Smigel, M. D. & Gilman, A. G. *J. biol. Chem.* **257**, 11416-11423 (1982).
14. Kurose, H., Katada, T., Haya, T., Ichijima, A. E. & Ui, M. *J. biol. Chem.* **261**, 6423-6428 (1986).
15. Gilman, A. G. *Rev. Biochem.* **56**, 615-649 (1987).
16. Sternweis, P. C., Northup, J. K., Smigel, M. D. & Gilman, A. G. *J. biol. Chem.* **256**, 11517-11526 (1981).
17. Ross, E. M., Howlett, A. C., Ferguson, K. M. & Gilman, A. G. *J. biol. Chem.* **253**, 6401-6412 (1978).
18. Florio, V. & Sternweis, P. C. *J. biol. Chem.* **260**, 3477-3483 (1985).
19. Berrie, C. P., Birdsall, N. J. M., Burgen, A. S. V. & Hulme, E. C. *Biochem. biophys. Res. Commun.* **87**, 1000-1006 (1979).
20. Ehler, F. J., Roeske, W. R. & Yamamura, H. I. *Fedn Proc.* **40**, 153-161 (1981).
21. Gurwitz, D., Kloog, Y. & Sokolovsky, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3650-3654 (1984).

22. Newman, M. E. & Belmaker, R. H. *Neuropharmacology* **26**, 211-217 (1987).
23. Seamon, K. B. & Daly, J. W. *Adv. Cyclic Nucleotide Res.* **20**, 1-150 (1986).
24. Batty, I. & Nahorski, R. *J. Neurochem.* **45**, 1514-1521 (1985).
25. Casebolt, T. L. & Jope, R. S. *Eur. J. Pharmac.* **140**, 245-246 (1987).
26. Abdel-Latif, A. A. *Pharmac. Rev.* **38**, 227-272 (1986).
27. Cockcroft, S. & Gomperts, B. D. *Nature* **314**, 534-536 (1985).
28. Litosch, I., Wallis, C. & Fzin, J. N. *J. biol. Chem.* **260**, 5464-5471 (1985).
29. Janowsky, D. S., El-Yousef, M. K., Davis, J. M. & Sekerke, M. J. *Lancet* **ii**, 632-635 (1972).

Characterization of the common genetic defect in humans deficient in debrisoquine metabolism

Frank J. Gonzalez*, Radek C. Skoda†, Shiko Kimura*, Morio Umeno*, Ulrich M. Zanger†, Daniel W. Nebert‡, Harry V. Gelboin*, James P. Hardwick§ & Urs A. Meyer†

* Laboratory of Molecular Carcinogenesis, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

† Department of Pharmacology, Biocenter, University of Basel, Basel CH-4056, Switzerland

‡ Laboratory of Developmental Pharmacology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

§ Division of Biological and Medical Research, Argonne National Laboratories, Argonne, Illinois 60517, USA

In population studies of individuals given the antihypertensive drug debrisoquine¹⁻³, two distinct phenotypes have been described: extensive metabolizers excrete 10-200 times more of the urinary metabolite 4-hydroxydebrisoquine than poor metabolizers. In family studies the poor-metabolizer phenotype behaves as an autosomal recessive trait with an incidence between 5% and 10% in the white population of Europe and North America¹⁻⁴, and extends to the deficient metabolism of more than 20 commonly prescribed drugs¹⁻⁴. Clinical studies have shown that such individuals are at high risk for the development of adverse side effects from these and probably many other drugs¹⁻⁶. Here we show that poor metabolizers have negligible amounts of the cytochrome P450 enzyme P450db1. We have cloned the human P450db1 complementary DNA and expressed it in mammalian cell culture. Furthermore, by directly cloning and sequencing cDNAs from several poor-metabolizer livers, we have identified three variant messenger RNAs that are products of mutant genes producing incorrectly spliced db1 pre-mRNA, providing a molecular explanation for one of man's most commonly defective genes (frequency of mutant alleles 35-43%).

Figure 1 shows that the antibody⁷ against rat P450db1 recognizes the human enzyme involved in the debrisoquine polymorphism. For these studies, we used the β -adrenergic blocking agent bufuralol as the prototype substrate for the human P450db1 enzyme⁶, because a highly sensitive assay for bufuralol 1'-hydroxylase activity has been developed⁸. Anti-(rat)P450db1 serum inhibits human liver microsomal bufuralol 1'-hydroxylase (Fig. 1a) and the human purified db1 in an *in vitro* reconstituted enzyme assay (Fig. 1b).

We measured the amount of human db1 protein in 29 liver samples using anti-(rat)db1 antibody on Western immunoblots (Fig. 1c, d). The microsomal bufuralol 1'-hydroxylase activity (Fig. 1d) was highly correlated with the concentration of immunodetectable human db1 protein (Fig. 1e). Using the previously prepared antibodies against human P450 with high bufuralol⁹ or debrisoquine¹⁰ hydroxylase activities, we found no correlation between the immunoreactive protein and enzymic activity; presumably, other proteins of identical molecular weight are recognized by these antibodies. The results shown in Fig. 1 indicate that the antibody to rat P450db1 can be used to isolate the human db1 cDNA.

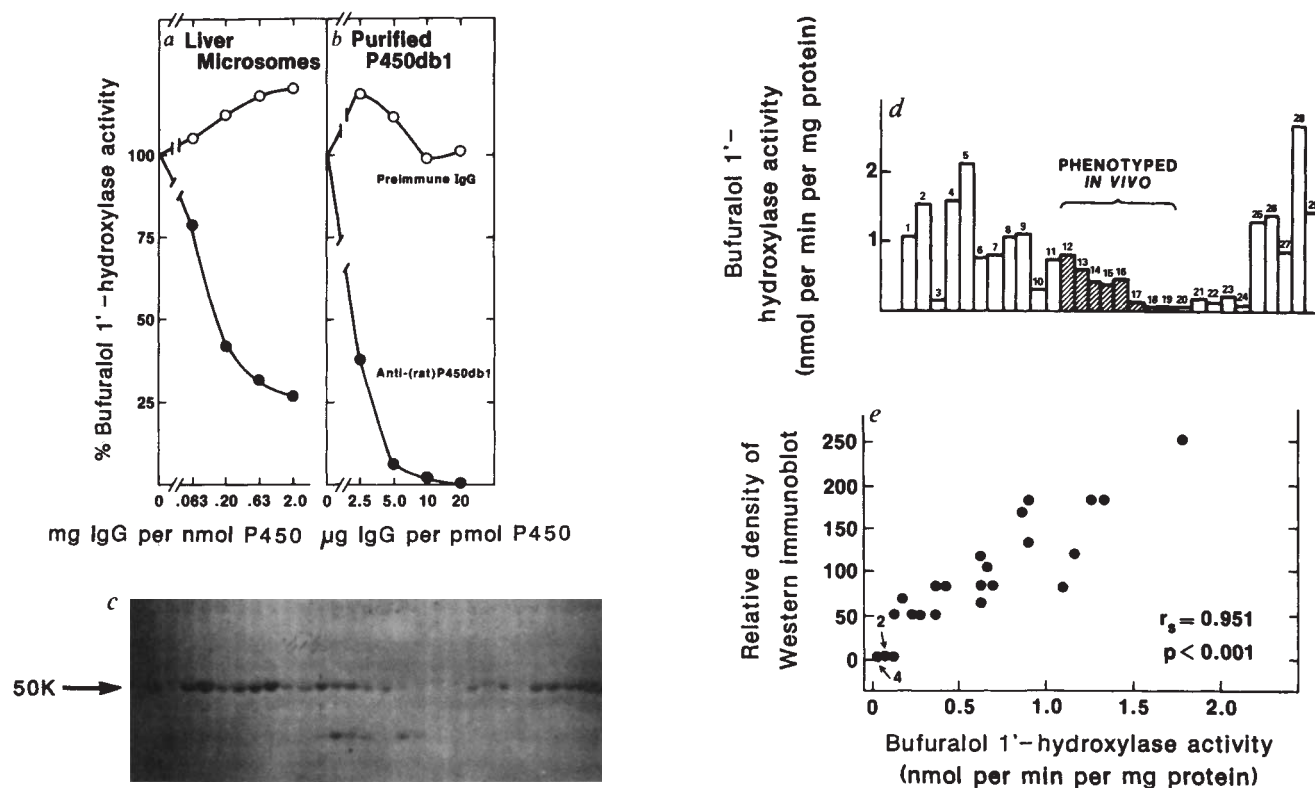


Fig. 1 Immunochemical correlation of anti-(rat)P450db1 with human liver bufuralol 1'-hydroxylase activity. *a*, Inhibition of the microsomal hydroxylase activity by antibody to rat db1. *b*, Inhibition of the purified db1 enzyme activity by anti-(rat)db1 antibody. *c*, Western immunoblot analysis of 29 human liver microsomal samples with anti-(rat)db1 antibody. *d*, Histogram of bufuralol 1'-hydroxylase activity in 29 human liver microsomal samples. *e*, Correlation of the immunochemically detectable protein with enzyme activity.

Methods. *a*, Increasing amounts of antibody (IgG) or preimmune IgG were incubated with human liver microsomes at 4 °C for 30 min before assaying for bufuralol 1'-hydroxylase activity⁸. The 100% of original hydroxylase activity represents 0.43 nmol min⁻¹ mg⁻¹ microsomal protein from an EM individual. *b*, Human db1 was purified as previously described⁹. Preincubation of the antibody or preimmune serum was carried out as above. The 100% of original hydroxylase activity represents 12.9 nmol of product formed per min per nmol P450 in the reconstituted assay. *c*, Electrophoretic separation of microsomal proteins (50 µg), electrotransfer to nitrocellulose filters, and immunoblot analysis with anti-(rat)db1 antibody (20 µg IgG ml⁻¹) were carried out as previously described^{7,9}. *d*, The eight individuals supplying specimens 12–19 were first phenotyped *in vivo* by quantitation of the urinary metabolites of either sparteine³ or debrisoquine¹ before microsomal activity was determined. The other samples, supplied by organ transplant donors, were assayed only *in vitro*; this protocol had been approved in advance by the hospital ethics committee. *e*, The intensity of the Western immunoblot db1 bands (Fig. 1c) was measured with a CAMAG TLC Scanner in the reflection mode, and the results were compared with the respective bufuralol 1'-hydroxylase activities of 27 individuals by means of the Spearman rank-correlation (two samples were not included because of poor resolution on the Western immunoblots).

Using a human λ gt11 expression library constructed from poly(A)⁺ mRNA (refs 11, 12) from the liver of an extensive metabolizer (EM), we isolated a 1,568-base-pair cDNA clone by screening with anti-(rat)db1 antibody. Analysis of the deduced amino acid sequence verified that this cDNA is one of a distinct subfamily of the P450II gene family¹³. The db1 cDNA probe was used to rescreen two other human cDNA libraries, each constructed from an EM liver. Several cDNAs were isolated, partially sequenced, and found to be identical to db1, suggesting that—in contrast to the rat⁷—the only active hepatic P450 gene in the human IID subfamily is db1.

To prove that the db1 cDNA codes for the appropriate P450 enzyme having bufuralol 1'-hydroxylase activity, we inserted the cDNA into the SV40- and adenovirus-based¹⁴ expression vector p91023(B) and transfected the construct into COS cells. Western immunoblot analysis of total extract from transfected cells (Fig. 2) revealed a protein that comigrated with human liver microsomal db1 protein; these cells expressed high bufuralol 1'-hydroxylase activity. The db1 protein was absent in cells transfected with the vector alone, and these cells contained no detectable bufuralol 1'-hydroxylase activity. These results confirm the identity of the db1 cDNA.

To study the molecular basis of the debrisoquine polymorphism, we performed Northern blot analysis on mRNA from seven livers studied in Fig. 1, including three poor-metabolizer (PM) livers with markedly decreased bufuralol 1'-hydroxylase activity (Fig. 3; lanes 2, 5 & 7). In all four EM livers exhibiting high bufuralol 1'-hydroxylase activity (Fig. 4; lanes 1, 3, 4 & 6), the probe hybridized to a mRNA of ~1.8 kilobases (kb). Variant mRNAs a and b (arrowed) are larger than, variant mRNA b' is considerably smaller than, and variant mRNA c is of similar size to the db1 wild-type (wt) mRNA. If the amount of the RNA detected with the db1 probe is normalized to the level of human actin mRNA, the quantity of the 2.1-kb variant a mRNA (lanes 2 & 5) is similar to that of the 2.1-kb variant b mRNA (lane 7). The weak hybridization signal in the wt db1 mRNA region of lanes 5 and 7 is ~1,000-fold less intense than the wt mRNA in EM livers and is probably due to cross-hybridization of the probe to other mRNAs in the P450II family¹³. Increasing the filter hybridization or washing stringency totally abolishes these small hybridization signals. A probe derived from the 3'-nontranslated region of the db1 mRNA also detects no mRNA in lanes 5 and 7 of the same size as the normal wt db1 mRNA. The Fig. 3 data thus suggest that the PM samples in lanes 2 and

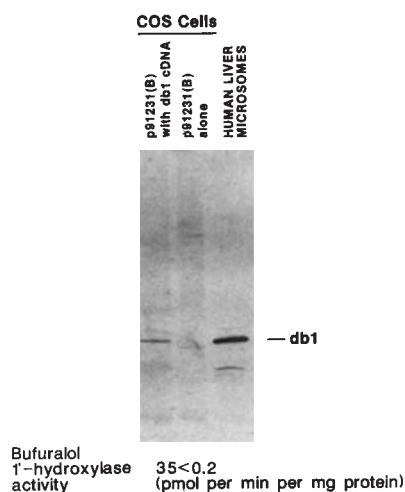


Fig. 2 Western immunoblot analysis of monkey COS cells transfected with p91231(B) containing the human db1 cDNA or with the p91231(B) vector alone. The db1 cDNA included 12 bases upstream from the initiation codon to 44 bases downstream from the termination codon. The cells were grown in 175-cm² flasks until ~50% confluent and transfected with 25 µg of plasmid DNA as described¹⁹. Cell extracts were prepared by sonication 72 h after transfection, and the protein (50 and 60 µg in lanes 1 and 2, respectively) was assayed by Western blot analysis^{7,9} and for bufuralol 1'-hydroxylase activity⁸. Human microsomal protein (50 µg) from an EM liver was used as a positive control.

7 possess at least two RNA species related to db1 mRNA, whereas the PM sample in lane 5 possesses at least one such RNA species.

To determine the nature of these mRNAs, we constructed λgt11 cDNA libraries and screened them with the db1 cDNA probe. Two classes of db1 cDNA from liver 2, one class from liver 5, and two classes from liver 7 were isolated from the cDNA libraries. At least 15 cDNAs from each library were characterized.

We found that variant a from both livers 2 and 5 retained intron 5 (Fig. 4), variant b from liver 7 retained intron 6, and variant b' from liver 7 lost the 3' half of exon 6 in combination with the removal of intron 6. That these clones represent nuclear RNA precursors is unlikely, because at least four individual cDNA clones of each of the b and b' variants, three individual cDNA clones of variant a from the liver 5 library, and three individual cDNA clones of variant a from the liver 2 library were partially sequenced across their respective variant regions and found to be identical. Further, the fact that variant mRNAs a and b are larger and variant mRNA b' is smaller than the wt db1 mRNA (Fig. 3) is consistent with the aberrant splicing defects observed by sequencing (Fig. 4).

Sequencing of cDNA for variant c from liver 2 revealed that the protein was identical to that of the wt db1 protein, except for several amino acid substitutions (unpublished data). Liver 2 had a very low but detectable amount of db1 protein and very low bufuralol 1'-hydroxylase activity, but much of this residual activity could be inhibited by anti-(rat)db1 antibody—in contrast to the lack of inhibition of residual bufuralol 1'-hydroxylase activity in liver microsomal samples 5 and 7 (in which no db1 protein is detectable). These data suggest that liver 2 has one mutant allele that gives rise to variant a mRNA, from which normal P450 protein is not produced, and another allele producing a normal-sized mRNA that is presumably translated into a partially altered and/or unstable protein. This hypothesis is currently under investigation.

Although full-length cDNAs for variants a and b were not obtained, we believe that both mRNAs contain the complete 5'

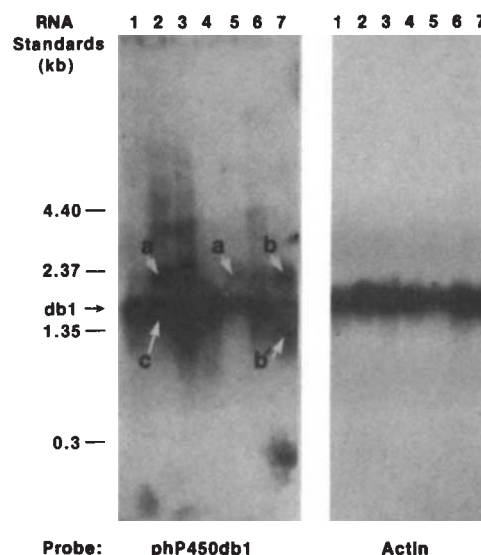


Fig. 3 Northern hybridization analysis of RNA from the livers of four EM and three PM individuals. Each patient was first phenotyped by the bufuralol 1'-hydroxylase assay⁸ of liver microsomes *in vitro*. The bufuralol 1'-hydroxylase activity in nmole per min per mg protein of the samples was; lane 1, 0.90 (EM), lane 2, 0.033 (PM); lane 3, 1.3 (EM); lane 4, 1.8 (EM); lane 5, 0.11 (PM); lane 6, 0.66 (EM); lane 7, 0.046 (PM). Another portion of each tissue was used for isolating poly(A)⁺-enriched RNA (ref. 20). These 7 patients are included in the 29 individuals examined in Fig. 1. Each RNA sample (0.1 µg) was electrophoresed on 1% agarose-2.2 M formaldehyde gels²¹ and transferred to NytranTM filters. The filters were hybridized with ³²P-labelled nick-translated phP450db1 and (human) actin probes under conditions previously described²². Autoradiography was carried out for 24 h at -80 °C with the aid of a Dupont Lightning-Plus intensifying screen. The three PM samples in lanes 2, 5 and 7 correspond to livers 20, 22 and 24 in Fig. 1b. Arrows a, b, b' and c represent the four mRNA variants that have been partially sequenced following cDNA cloning. RNA standards (Bethesda Research Laboratories) were included and identified by hybridization of the filters with nick-translated λ DNA.

end and thus are transcribed from the normal initiation start site. This conclusion is based on the lengths of the RNAs and the fact that these RNAs hybridize to a 300-base pair (bp) probe (the 5'-most 300 bases) derived from the normal db1 mRNA. In contrast, variant b' mRNA is ~1.0 kb long (Fig. 3) and does not hybridize with this 300-bp 5' db1 probe. It is therefore likely that the transcription of variant b' is initiated from an artificial promoter region that formed after the original gene-inactivating base change had taken place. Because the cDNA sequences of variants b and b' are identical in their common coding and 3'-noncoding regions, it seems quite likely that the retention of intron 6 in variant b and the aberrant splice (use of a cryptic promoter) in variant b' arise from the same mutation(s). The possibility that both variant b and b' mRNAs are derived from the same allele cannot, therefore, be excluded.

Both variant mRNAs a and b contain the normal GT/AG consensus exon-intron junction signal, suggesting that base change(s) in another, less obvious, spliceosome recognition region might cause the mutation in these db1 genes. If we had directly sequenced these mutant genes, the observed defects (intron retention; partial exon removal; initiation from cryptic promoter) would not have been readily apparent. Sequencing of the entire normal human P450db1 gene (introns, exons and flanking regions) will be necessary before the base mutations that cause the aberrant splicing defects in these mutant P450db1 genes can be fully understood. It is well known that abnormal mRNA often leads to an unstable protein product¹⁵.

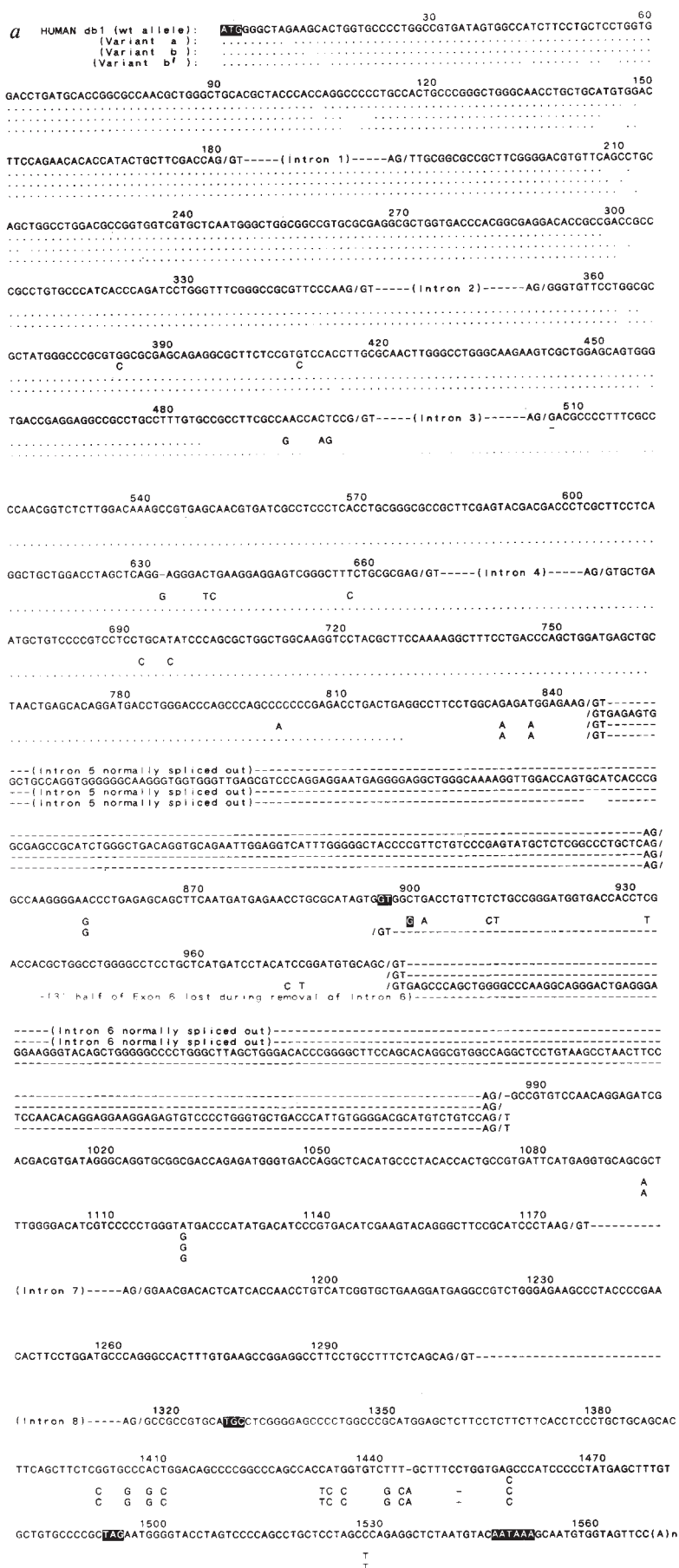


Fig. 4 Top, nucleotide sequence of the human wild-type (wt) db1 cDNA and of the cDNAs for variants a, b and b'. The intron-exon junctions were assigned on the basis of the partial sequencing of a human P450db1 genomic clone. The 5' portions of the variant cDNAs not sequenced are denoted by dots. The wt sequence is numbered; the variant sequences are shown only at those positions where there is no correlation. The initiation codon, possible base(s) involved in aberrant splicing, codon for cysteine that participates in the haem-binding site¹³, termination codon, and the putative poly(A) addition signal are enclosed in blackened boxes. Bottom, diagram of the P450db1 exons and approximate sizes of the eight introns of the db1 gene. The normal splicing pattern of the wt pre-mRNA, and aberrant splicing patterns of the pre-mRNA from variants a, b and b', are shown. To explain the size of the b' mRNA (Fig. 3), we propose that the initiation of transcription is probably at a cryptic promoter within intron 3 (see text). The normal db1 protein (497 amino acids) would differ from the prematurely terminated, truncated proteins of all three variants. This probably results in unstable proteins that are so rapidly degraded that the anti-(rat)db1 antibody cannot detect them. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00300.

Our experiments establish that the basis of the 'poor metabolism' phenotype is, in most cases examined, an absence of detectable human liver microsomal P450db1 protein. How many other mutant alleles can be responsible for the PM phenotype remains to be determined.

There have been reports suggesting an association between the PM or EM phenotype and human disease¹⁶, including cancer^{17,18}. The significance of these cancer studies is unclear, as the db1 enzyme is not known to metabolize well-established environmental carcinogens. The enzyme may, however, metabolize other, as yet undescribed, dietary carcinogens, or the db1 gene may be linked to an oncogene, tumour suppressor gene, or genes encoding enzymes that do metabolize carcinogens.

It is particularly noteworthy that the PM phenotype is deficient in the metabolism of debrisoquine and bufuralol, as well as being unable to metabolize a growing list of ~24 drugs, including many other β -adrenergic blocking agents, antiarrhythmics, antidepressants, the hallucinogen 4-methoxyamphetamine and commonly used drugs such as the antitussive opioid dextromethorphan¹⁻⁴. Identification of the PM individual before administering any of these drugs would be of great value to the physician. The successful expression of the db1 cDNA-encoded enzyme activity (Fig. 2) in COS cells, with no detectable background of P450-mediated drug metabolism, could lead to an assay in which new drugs being developed can be tested to see whether they suffer from this particular drug oxidation polymorphism.

We thank our colleagues for a critical reading of this manuscript. We also thank P. Beaune (Paris), M. Eichelbaum

(Stuttgart), W. Kalow and T. Inaba (Toronto) for supplying human liver samples and Ingrid E. Jordan for expert secretarial assistance. Supported by the Swiss National Research Foundation (U.A.M.).

Received 29 September; accepted 11 December 1987.

1. Idle, J. R. & Smith, R. L. *Drug Metab. Rev.* **9**, 301-317 (1979).
2. Kupfer, A. & Preisig, R. *Semin. Liver Dis.* **3**, 341-354 (1983).
3. Eichelbaum, M. *Fed. Proc.* **43**, 2298-2302 (1984).
4. Price-Evans, D. A. in *Ethnic Differences in Reactions to Drugs and Xenobiotics* (eds Kalow, W., Goedde, H. W. & Agarwal, D. P.) 491-526 (Liss, New York, 1986).
5. Davies, D. S., Kahn, G. C., Murray, S., Brodie, M. J. & Boobis, A. R. *Br. J. clin. Pharmacol.* **11**, 88-91 (1981).
6. Meyer, U. A., Gut, J., Kronbach, T., Skoda, C., Meier, U. T. & Catin, T. *Xenobiotica* **16**, 449-464 (1986).
7. Gonzalez, F. J. *et al. DNA* **6**, 149-161 (1987).
8. Kronbach, T., Mathys, D., Gut, J., Catin, T. & Meyer, U. A. *Analyt. Biochem.* **162**, 24-32 (1987).
9. Gut, J., Catin, T., Dayer, P., Kronbach, T., Zanger, U. & Meyer, U. A. *J. biol. Chem.* **261**, 11734-11743 (1986).
10. Distlerath, L. M. *et al. J. biol. Chem.* **260**, 9057-9067 (1985).
11. Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1194-1198 (1983).
12. Watson, C. J. & Jackson, J. F. in *DNA Cloning*, Vol. 1 (ed. Glover, D. M.) 79-88 (IRL, McLean, 1985).
13. Nebert, D. W. & Gonzalez, F. J. *A. Rev. Biochem.* **56**, 945-993 (1987).
14. Wong, G. G. *et al. Science* **228**, 810-815 (1985).
15. Orkin, S. H. & Kazanian, H. H. Jr. *A. Rev. Genet.* **18**, 131-171 (1984).
16. Baer, A. N., McAllister, C. B., Wilkinson, G. R., Woosley, R. L. & Pincus, T. *Arthritis Rheum.* **29**, 843-850 (1986).
17. Ayesh, R., Idle, J. R., Ritchie, J. C., Crothers, M. J. & Hetzel, M. R. *Nature* **312**, 169-170 (1984).
18. Idle, J. R. *et al. Cancer Lett.* **11**, 331-338 (1981).
19. Luthman, H. & Magnusson, G. *Nucleic Acids Res.* **11**, 1295-1308 (1983).
20. Chirgwin, J. M., Przybyla, A. W., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
21. Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. *Biochemistry* **16**, 4743-4751 (1977).
22. Church, G. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991-1995 (1984).

A new human immunoglobulin V_H family preferentially rearranged in immature B-cell tumours

C. G. Humphries, A. Shen*, W. A. Kuziel, J. D. Capra, F. R. Blattner* & P. W. Tucker

Department of Microbiology, University of Texas Health Science Center, Dallas, Texas 75235, USA

* Department of Genetics, University of Wisconsin, Madison, Wisconsin 53706, USA

The prevalent forms of adult and childhood B-cell neoplasia are chronic lymphocytic (CLL) and acute lymphocytic (ALL) leukaemia, and are typified by a nearly monoclonal accumulation of cells expressing a single heavy (H) and light (L) chain variable (V) region. V gene selection could be random, or quite biased if the disease or the developmental status of the transformed cell somehow influenced DNA rearrangement. We have cloned and sequenced three germ-line V_H gene segments that constitute a new human V_H family (subgroup V) linked within 160 kilobase pairs of the D_H-J_H complex. One V_H(V) member is rearranged in about 30% of patients with CLL and ALL, but not in IgM-expressing B-cell lines from peripheral blood. In some tumours, we detect a truncated (V_H(V) RNA devoid of constant regions that originates from unrearranged V_H(V) genes. In other tumours and in resting splenocytes, we detect large amounts of normally sized V_H(V)-associated mRNA, although stimulation by mitogen of splenic B cells results in loss of V_H(V)-hybridizing RNA. These features suggest that biased rearrangement of subgroup V may be under developmental selection.

Immunoglobulin V_H segments can be divided into families according to amino-acid and nucleic-acid similarities. In mice, there are a minimum of nine non-cross-hybridizing (<70-80%

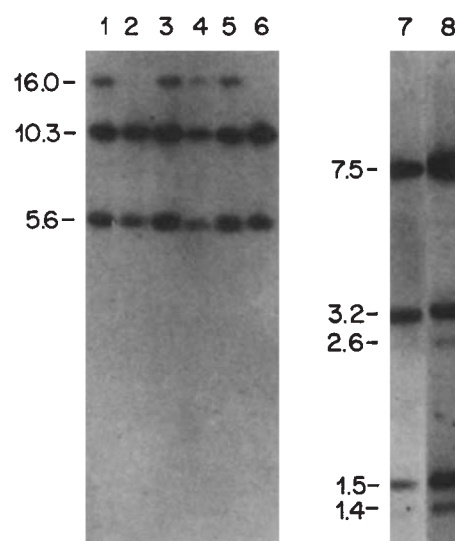


Fig. 1 Southern blot hybridization of the human V_H subgroup V family. Lanes 1-6, DNAs prepared from six normal livers were digested with *Hind*III and analysed under high-stringency wash conditions (0.1 × SSC, 60 °C). Lanes 7 and 8, DNA from the donor of lane 1, cut with *Hind*III and *Bam*HI and analysed at high (lane 7) and low (0.3 × SSC, 42 °C) stringency. Sizes (in kbp) are indicated.

Methods. DNA was prepared as described²⁹, digested with fivefold excess of restriction enzyme and fractionated through 0.8% agarose. DNA was transferred to nitrocellulose filters and hybridized for 24-36 h to ³²P-labelled 900-base pair (bp) *Xba*I-*Stu*I fragment carrying the V_H251 gene and its 5' flanking region at 42 °C in 50% formamide, 5 × SSC (1 × SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.02% polyvinylpyrrolidone, 0.02% Ficoll and 100 μg ml⁻¹ denatured herring sperm DNA. Filters were washed in 3 × SSC, 0.01% SDS at room temperature, then at the stringencies indicated above, air-dried and exposed to X-ray film at -70 °C using an intensifying screen.

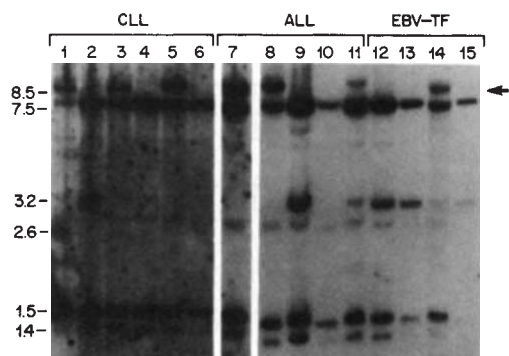


Fig. 3 Biased rearrangement of $V_H(V)$ member V_H251 in chronic and acute lymphocytic leukaemia DNA. Lanes 1, 3 and 5 show examples of CLL DNA rearrangement of V_H251 *Bam*HI-*Hind*III-containing fragment from 7.5 to 8.5 kb (marked on the right with arrow). All samples (lanes 1-6) were obtained by leukaphoresis of patients in active leukaemic phase. Lanes 7 and 8, established pre-B ALL lines, NALM8 and NALM16, respectively³⁷; lanes 9-11, freshly isolated ALL samples. Lane 14, EBV-transformed pre-B cell line FB8-5A10, prepared from the bone marrow of a 14-week normal fetus. Lanes 12, 13, 15, EBV-transformed, IgM-secreting cell lines prepared from normal adult peripheral blood lymphocytes. Sizes (in kbp) are indicated to the left.

Methods. DNA preparations and blotting are described in Fig. 1 legend. Probability (P) that a single gene among 200 rearranged by chance is

$$P = \frac{n!}{x!(n-x)!} \left(\frac{1}{200} \right)^x \left(\frac{199}{200} \right)^{n-x}$$

where the probability of V_H251 rearrangement is 1 in 200; the probability of a non- V_H251 rearrangement is 199/200; x is the number of V_H251 rearrangements and n is the total of rearrangements analysed. Calculation of probability according to binomial expansion assumes all trials (rearrangements) are independent.

and 1.4 kb fragments, and the V_H gene segment cloned on the 16 kb fragment (V_H32) reduces to 3.2 kb (Fig. 1, lane 7). Analysis of various liver DNAs (Fig. 1, lanes 1-6) showed that V_H32 was absent either on both (lanes 2 and 6) or one (lane 4) allele.

Figure 2 presents an alignment of V_H251 and the related genes. As anticipated, V_H251 showed most homology with the CLL-derived probes (95 and 88% overall)¹⁰. V_H251 was also the only one of the set that contained an amino-acid open-reading frame. It has all the features^{3,11} characteristic of a functional V gene. V_H32 and V_H15 were highly related to V_H251 (95 and 86%, respectively, from leader to 3' recombination nonamer). V_H32 and V_H15 shared 85% identity. No published amino-acid or DNA sequences³ show similarities beyond 64% to V_H251 , V_H32 or V_H15 . Therefore, this set of three genes constitute a new human V_H family, subgroup V. The closest relative, V_H20 , belongs to subgroup I because of its 87% relatedness to another member⁵. Unlike the four human V_H subgroups previously analysed, $V_H(V)$ does not cross-hybridize to any of the nine murine V_H families (data not shown). V_H105 , physically linked to V_H251 , was only 56% similar, whereas the weakly hybridizing V_H20 was 67% similar to V_H251 . V_H105 is probably the allelic counterpart (97% similar) of the subgroup II sequence, V_H11 . Linkage of subgroup II (V_H105) and V (V_H251) genes is consistent with previous reports of interspersal of unrelated V_H families in the human^{4,7}. The 3' ends of V_H105 and V_H15 are highly related (denoted by the dots in Fig. 2), yet these two genes are only 52% similar in the coding regions, suggesting that some type of sequence exchange has occurred.

The finding that V_H251 was rearranged in two cases of familial CLL (ref. 10) prompted a broader analysis of tumours and EBV-transformed 'normal' B cells. Using a $V_H(251)$ fragment as a probe, we analysed DNA from peripheral blood mono-

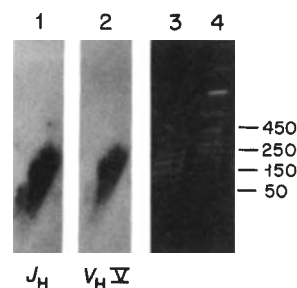


Fig. 4 Linkage of $V_H(V)$ to J_H using pulsed-field electrophoresis. Genomic DNA from the human HeLa carcinoma cell line was digested with *Cla*I. The same filter was hybridized with a J_H probe (lane 1), stripped and reprobed with a $V_H(V)$ probe (lane 2). Markers are phage λ concatamers (lane 3) and yeast chromosomes (lane 4) derived from AB972.

Methods. Cell nuclei were prepared as described³⁸ and embedded in agarose plugs³⁹. A 1% agarose gel was run in 0.5 $\times$ TBE at 250 v (ref. 40) using 45-s pulse intervals. The $V_H(V)$ probe is described in Fig. 1 legend; the J_H probe was a 1kb *Sau*3A fragment spanning J_H2 - J_H6 (ref. 41).

nuclear cells isolated during the leukaemic phase of the disease (Fig. 3, lanes 1-6). In 9 out of 33 CLLs, we detected a rearranged *Hind*III-*Bam*HI fragment of 8.5 kb (lanes 1, 3 and 5), the same size as the V_H251 rearrangement in the familial CLLs (ref. 10). A V_H251 -specific probe derived from the 5' flank also detected the 8.5 kb band (data not shown), confirming that V_H251 is the only $V_H(V)$ member rearranged. We then isolated DNA from ALL samples, comprising established cell lines and isolates from leukaemic peripheral blood or bone marrow. Six of the 16 DNAs tested contained the 8.5 kb V_H251 rearrangement (lanes 8, 11). No rearrangement of V_H251 was observed in 38 human B-cell lines derived by clonal transformation with EBV (lanes 12, 13, 15) with the single exception (lane 14) of a bone-marrow-derived line of pre-B origin. Therefore, rearrangement of V_H251 , but not of other $V_H(V)$ member genes, appears to be biased only in tumours or transformed cells representing early stages of B-cell maturation.

The mouse V_H 7183 family is preferentially rearranged in pre-B cells² and maps closest to the D_H - J_H cluster^{1,12}, where J is the joining segment, D the diversity segment. We therefore examined $V_H(V)$ linkage by pulsed-field electrophoresis (Fig. 4). Using five restriction enzymes, there was a concordance of hybridization of J_H and $V_H(V)$, ranging from the smallest (158 kb with *Cla*I, Fig. 4) to the largest (350 kb with *Mlu*I). $V_H(I)$ and $V_H(II)$ probes hybridized to the same restriction fragments as $V_H(V)$ and J_H , but these probes also hybridized to other larger fragments. This indicates that the $V_H(V)$ family is relatively close to D_H and J_H , but that other V_H families are interspersed within the cluster. The data are consistent, but not conclusively so, with the interpretation that $V_H(V)$ is the most D_H - J_H proximal family in man.

Several of the tumours that rearrange V_H251 were analysed for steady-state levels of $V_H(V)$ -hybridizable mRNA (for example, Fig. 5, lane 1) and full-length μ -chain mRNA is usually observed (lane 2). Certain ALLs and CLLs (lane 3) transcribe a truncated, (~1.2 kb) polyadenylated RNA that hybridizes to $V_H(V)$ and is devoid of constant-region sequences. To date, we have detected the truncated species in three tumours. In each case the $V_H(V)$ genes are not rearranged, but full-length μ -mRNA is transcribed from the other allele (as in lane 7). The 1.2 kb RNA hybridizes to a probe generated 3' to the V_H251 -coding region (data not shown), suggesting that it is a product of a germline V_H251 gene. This is reminiscent of the V_H transcripts from unrearranged J558 V_H genes in murine pre-B cells and fetal liver¹³. Analysis of RNA isolated from normal adult resting splenocytes showed $V_H(V)$ -containing RNA species

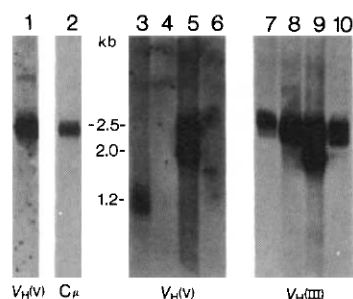


Fig. 5 Steady-state levels of $V_H(V)$ -associated transcripts in leukaemic and normal cells. Probes are indicated below each panel and hybridizing transcript sizes (in kilobases) are shown. RNA sources include: tumour cells from a patient with CLL (lanes 1, 2), corresponding to the DNA in Fig. 3, lane 3; a human ALL (lanes 3, 7); the human CLL line SED (lanes 4, 8); splenocytes from a 9-month-old normal transplant donor (lanes 5, 9); a fraction of the same splenocytes cultured for 3d in RPMI 1640 culture medium with 10% fetal calf serum, penicillin (100 units per ml), streptomycin (100 $\mu\text{g ml}^{-1}$), 2mM L-Glutamine, 1mM sodium pyruvate, 0.1mM non-essential amino acids in the presence of formalinized *Staphylococcus aureus* (Calbiochem-Behring) at a final concentration of 1/60,000 (v/v) (lanes 6 and 10).

Methods. Total RNA was isolated by the guanidinium thiocyanate/cesium chloride spin method⁴². Total RNA (10 μg , lanes 1 and 2) or poly(A)⁺ RNA (10 μg in lanes 4, 6, 8, 10 or 20 μg in lanes 3, 5, 7, 9) was fractionated by formaldehyde/agarose gel electrophoresis, transferred to Gene Screen nylon membranes, and hybridized to gel-purified DNA fragments as described previously⁴³. The same filters, comprising lanes 1 and 2 or lanes 3–6 and 7–10, were stripped by washing in water at 100°C and used for both the hybridizations shown. Autoradiography was performed at –70°C with intensifying screens for 5 d (lanes 1, 2), 7 d (lanes 3–6) or 1 d (lanes 7–10). The $V_H(V)$ probe is described in Fig. 1 legend; the $V_H(III)$ probe is a 600-bp *Sau3A-HhaI* fragment derived from the genomic DNA clone pVHB.26 (ref. 4); the C_μ probe is a 1.2kb *EcoRI* fragment which includes the $C_\mu 1$, $C_\mu 2$ and $C_\mu 3$ exons⁴⁴.

(lane 5) whose sizes are consistent with full-size μ -, δ - and γ -transcripts (W.A.K. *et al.*, manuscript in preparation). But it is difficult to detect $V_H(V)$ -containing mRNA in the same splenocyte preparation after stimulation with a B-cell mitogen (lane 6), even though full-size $V_H(III)$ -containing μ -transcripts are abundant (lane 10). Therefore, immature (IgM/IgD-producing) and apparently class-switched (IgG-producing) splenic B cells express $V_H(V)$, but most of the proliferating B cells in the mitogen-activated spleen do not.

Based on an estimate of 200 human V_H genes (J. Berman and F. Alt, personal communication), the probability (Fig. 3 legend) that 9 out of 33 V_H251 rearrangements in CLL occurred by chance would be 5×10^{-11} ; similarly, chance occurrence of 6 out of 16 V_H251 rearrangements in ALLs is 1×10^{-8} . Thus, a non-stochastic process is indicated. Antigen-independent, developmental regulation of V_H gene usage observed in some mouse strains has been correlated with rearrangement of the more D_H - J_H proximal V_H families^{2,12,14–16}. Biased rearrangement of V_H251 could therefore reflect the potentially D_H -proximal position of this gene (<160 kb). An alternative hypothesis is that biased rearrangement is selected by antigen. Possibilities include common leukaemogenic environment agents^{17,18}, or viruses transmitted vertically or horizontally. As postulated for the relationship of HTLV-I and CLL (ref. 19), leukaemia could result from the clonal stimulation of V_H251 -bearing B cells by either a neoplastic T cell or through virus that recognizes the V_H251 -associated idiotype.

CLL cells represent a unique B-cell subset in the human in that they express the Leu-1 antigen normally found on T cells²⁰. The normal Leu-1⁺ B-cell subpopulation²¹ and the equivalent subpopulation in mice (Ly-1⁺B) (refs. 21, 22) have been implicated in auto-regulatory and autoimmune functions. Indeed, human²³ and murine²⁴ rheumatoid factors, and other auto-

antibodies, arise from a restricted set of variable region genes. It has been shown that an anti-rheumatoid factor antibody which is specific for a member of the V_k III light-chain variable region family, reacts with the leukaemic cells of 5 out of 20 CLL patients²⁵. A second correlation can be drawn with the Leu-1⁺ subset and our observation that $V_H(V)$ splenic RNA levels are dramatically reduced after mitogen stimulation. Neither CLL cells²⁶ nor normal Ly-1⁺ (ref. 27) and fetal Leu-1⁺ B cells²⁸ proliferate or produce specific antibodies in response to lectins. Normal adults have few Leu-1⁺ B cells (generally <5% splenocytes), but they are a major population (~50%) in fetal spleen²⁸. CLL cells that consistently express the Leu-1 antigen could represent the neoplastic counterpart of this stage of B-cell differentiation. Similarly, $V_H(V)$ expression could identify a separate lineage of B cells in humans.

We thank Dr Graham Smith and Dr William Blattner for CLL and ALL isolates, and Drs Richard Wasserman, Peter Stasney, Maria Hoover, Max Cooper and Carol Webb for EBV-transformed lines. We thank Chhaya Das and Carol Williams for technical assistance and Marilyn Gardner for preparation of the manuscript. This research was supported by the NIH.

Note added in proof: Since submission, Schroeder *et al.*⁴⁵ have described a sixth human V_H family. We and others (F. Alt, personal communication) have pulse-field mapped $V_H(VI)$ to within 70 kb of J_H on a fragment that does not contain $V_H(V)$. Thus, $V_H(V)$ appears not to be the most proximal family.

Received 14 September; accepted 18 December 1987.

1. Riblet, R. & Brodeur, P. H. in *Genetics and Molecular Immunology* (eds Herzenberg, L. A., Blackwell, C. & Herzenberg, L. A.) 89.1–89.6 (Blackwell, Oxford, 1986).
2. Yancopoulos, G. D. *et al.* *Nature* **311**, 727–733 (1984).
3. Kabat, E. A. *et al.* *Sequences of Proteins of Immunological Interest* Vol. 4 (U.S. Dept. of Health and Human Services, 1987).
4. Kodaira, M. *et al.* *J. molec. Biol.* **190**, 529–541 (1986).
5. Rechavi, G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 855–859 (1983).
6. Rechavi, G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 4405–4409 (1982).
7. Matthysens, G. & Rabbitts, T. H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6561–6565 (1980).
8. Lee, K. H., Matsuda, F., Kinashi, T., Kodaira, M. & Honjo, T. *J. molec. Biol.* **195**, 761–768 (1987).
9. Blattner, W. *et al.* *Ann. intern. Med.* **84**, 554–557 (1976).
10. Shen, A., Humphries, C., Tucker, P. W. & Blattner, F. R. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8563–8567 (1987).
11. Parslow, T. G., Blair, D. L., Murphy, N. J. & Granner, D. K. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2650–2654 (1984).
12. Rathbun, G. A., Capra, J. D. & Tucker, P. W. *EMBO J.* **10**, 2931–2937 (1987).
13. Yancopoulos, G. D. & Alt, F. W. *Cell* **40**, 271–281 (1985).
14. Wu, G. E. & Paige, C. J. *EMBO J.* **5**, 3475–3481 (1986).
15. Riley, S. C., Connors, S. J., Klinman, N. R. & Ogata, R. T. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2589–2592 (1986).
16. Perlmutter, R. M., Kearney, J. F., Chang, S. P. & Hood, L. G. *Science* **227**, 1597–1601 (1985).
17. Infante, P., Wagoner, J. & Rinskey, R. *Lancet* **ii**, 76–78 (1977).
18. Blair, A., Decoufle, P. & Grauman, D. *Am. J. Public Health* **69**, 508–511 (1979).
19. Mann, D. L. *et al.* *Science* **236**, 1103–1106 (1987).
20. Bousmell, L. H. *et al.* *J. exp. Med.* **152**, 229–239 (1980).
21. Casali, P., Inghirami, G., Nakamura, M., Davies, T. F. & Notkins, A. L. *Science* **234**, 476–480 (1986).
22. Hayakawa, K., Hardy, R. R., Parks, D. R. & Herzenberg, L. A. *J. exp. Med.* **157**, 202–207 (1983).
23. Manheimer-Lory, A., Momestier, M., Bellan, B., Alt, F. & Bona, C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8293–8297 (1986).
24. Kunkel, H. G., Agnello, V., Joslin, F. G., Winchester, R. J. & Capra, J. D. *J. exp. Med.* **137**, 331–342 (1979).
25. Kipps, T. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 2916–2920 (1986).
26. Fu, S. M., Chiorazzi, N. & Kunkel, H. G. *Immunol. Rev.* **48**, 23–35 (1979).
27. Hayakawa, K. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 2494–2498 (1984).
28. Antin, J. H., Emerson, S. G., Martin, P., Gadol, N. & Ault, K. A. *J. Immun.* **136**, 505–510 (1986).
29. Blin, N. & Stafford, D. W. *Nucleic Acids Res.* **3**, 2303–2308 (1976).
30. Liu, C.-P., Tucker, P. W., Mushinski, J. F. & Blattner, F. R. *Science* **209**, 1348–1352 (1980).
31. Loenen, W. & Blattner, F. R. *Gene* **26**, 171–179 (1983).
32. Tucker, P. W. in *Genetics and Molecular Immunology* (eds Herzenberg, L. A., Blackwell, C. & Herzenberg, L. A.) 86.1–86.16 (Blackwell, Oxford, 1986).
33. Viera, J. & Messing, J. *Gene* **19**, 771–780 (1982).
34. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
35. Bencini, D. A., O'Donovan, G. A. & Wild, J. R. *Biotechniques* **2**, 4–12 (1984).
36. Rubin, C. M. & Schmidt, C. W. *Nucleic Acids Res.* **8**, 4613–4622 (1980).
37. Hurwitz, R. *et al.* *Int. J. Cancer* **23**, 174–180 (1979).
38. Hieter, P. A., Hollis, G. F., Korsmeyer, S. J., Waldmann, T. A. & Leder, P. *Nature* **24**, 536–540 (1981).
39. Van der Ploeg, L. H., Schwartz, D. C., Cantor, C. R. & Borst, P. *Cell* **37**, 77–84 (1984).
40. Carle, G. F. & Olson, M. V. *Nucleic Acids Res.* **12**, 5647–5664 (1984).
41. Ravetch, J. *et al.* *Cell* **27**, 583–591 (1981).
42. Chirgwin, J. M., Przybyla, A. S., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5304 (1979).
43. Tutt, M. *et al.* *J. Immun.* **137**, 2998–3001 (1986).
44. Rabbitts, T. H., Forster, A. & Milstein, C. P. *Nucleic Acids Res.* **9**, 4509–4525 (1981).
45. Schroeder, H. W., Hilson, J. L. & Pearlmuter, R. M. *Science* **238**, 791–793 (1987).

Tracking kinesin-driven movements with nanometre-scale precision

Jeff Gelles*, Bruce J. Schnapp† & Michael P. Sheetz*

* Department of Cell Biology and Physiology, Washington University School of Medicine, Box 8101, 660 South Euclid Ave, St Louis, Missouri 63110, USA

† Laboratory of Neurobiology, National Institute of Neurological and Communicative Disorders and Stroke, National Institutes of Health at the Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA

Several enzyme complexes drive cellular movements by coupling free energy-liberating chemical reactions to the production of mechanical work¹⁻³. A key goal in the study of these systems is to characterize at the molecular level mechanical events associated with individual reaction steps in the catalytic cycles of single enzyme molecules. Ideally, one would like to measure movements driven by single (or a few) enzyme molecules with sufficient temporal resolution and spatial precision that these events can be directly observed. Kinesin, a force-generating ATPase involved in microtubule-based intracellular organelle transport⁴⁻¹⁰, will drive the unidirectional movement of microscopic plastic beads along microtubules *in vitro*^{4,9}. Under certain conditions, a few (≤ 10) kinesin molecules may be sufficient to drive either bead movement or organelle transport. Here we describe a method for determining precise positional information from light-microscope images. The method is applied to measure kinesin-driven bead movements *in vitro* with a precision of 1–2 nm. Our measurements reveal basic mechanical features of kinesin-driven movements along the microtubule lattice, and place significant constraints on possible molecular mechanisms of movement.

Analysis of the mechanisms of force-generating enzymes such as actomyosin, dynein and kinesin has been facilitated by the development of systems in which movement can be measured *in vitro* under controlled conditions^{4,11-13}. In the experimental system for kinesin^{4,12}, a suspension of plastic beads with kinesin nonspecifically adsorbed to their surfaces is applied to purified, taxol-stabilized microtubules adhering to a glass coverslip in a buffer solution containing ATP. Video-enhanced differential interference contrast (DIC) microscopy^{14,15} is used to observe the kinesin-coated beads attach to and move along the microtubules. The plastic beads used in this assay are smaller than the resolution limit of the optical microscope and therefore produce very low-contrast images. Our technique, designed to maximize the amount of positional information extracted from low-contrast images, measures bead position in individual video frames with a precision of 1–2 nm. Similar precision has been obtained previously in light-microscope position measurements of high-contrast images (for example, ref. 16). We recorded 30 images (frames) per second on a video-disk recorder, and analysed them to determine the x and y coordinates of the beads in each frame (Fig. 1).

To determine the precision of our technique, we recorded a video sequence of stationary beads immobilized on a glass coverslip. The measured positions of two beads in this test specimen, both taken from the same sequence of video frames, are plotted in Fig. 2*a, b*. In both cases, the points fall in a region of width ~ 34 nm and height ~ 6 nm and display a gradual positional drift with time during the period of the measurement. The direction and magnitude of the drift are the same in Fig. 2*a* and *b*, as is the overall shape of the spatial distribution of the data points. Such drift might be caused by small movements of the specimen relative to the microscope optics. To cancel its effect on the measured positions, we used one of the stationary beads as a fiducial reference point in the specimen and measured

Fig. 1 Stages in the high-precision measurement of relative bead position in a video frame. *a*, Image; *b*, kernel; *c*, cross-correlation of image and kernel; *d*, centroid calculation. The illustrated sequence of computations derives positional information from the entire bead image rather than from an individual point or edge only, and thereby maximizes the precision of the measurement. *a*, A segment of a digitized video frame containing the video-enhanced DIC microscope image of a bead. The DIC image of a single bead consists of apposed bright and dark areas. The segment consists of a 49×63 rectangular matrix of integers, $I(x, y)$, each of which represents the recorded light intensity at the point (x, y) in the image (x and y are integers). *b*, A 'kernel' segment $K(x, y)$, consisting entirely of a single bead image. The centre of this segment is taken to be the point $(0, 0)$. The kernel is used as a template or standard; the effect of the calculations below is to find the position in the frame segment shown in *a* for which the surrounding intensity distribution most closely matches that surrounding the centre of the kernel. When a sequence of consecutive video frames is analysed, a single kernel derived from one of the frames is used in the analysis of all of them. The x and y dimensions of the kernel are denoted α and β , respectively; for the case shown here, $\alpha = 27$ pixels and $\beta = 33$ pixels. *c*, Cross-correlation $C(x, y)$ of the frame segment *a* with a scaled version of the kernel *b*. The cross-correlation calculation

$$C(x, y) = \sum_{i=-\alpha/2}^{\alpha/2} \sum_{j=-\beta/2}^{\beta/2} I(x+i, y+j) \{K(i, j) - s\}$$

yields sharp peaks in regions where the intensity distribution in the frame segment closely matches that in the kernel^{32,33}. The scale term s is set to the mean intensity of the pixels in the kernel. *d*, Calculation of the centroid of the peak in the cross-correlation *c*. A threshold value T is subtracted from each point in the cross-correlation and points where the difference is negative are discarded; the figure shows the 219 remaining points. From these points the coordinates of the centroid (x_c, y_c) are computed

$$x_c = \frac{\sum x \{C(x, y) - T\}}{\sum \{C(x, y) - T\}} \quad y_c = \frac{\sum y \{C(x, y) - T\}}{\sum \{C(x, y) - T\}}$$

The point (x_c, y_c) is taken as the position of the bead in *a*. In a time sequence of frame segments like *a*, the set of calculated coordinates (x_c, y_c) accurately reflect the translational movements of the bead image providing that this image (and the peak in *d*) do not significantly change shape. The absence of such a shape change was verified for all reported data. (The z -axis scale in *c* and *d* is different from that in *a* and *b*.)

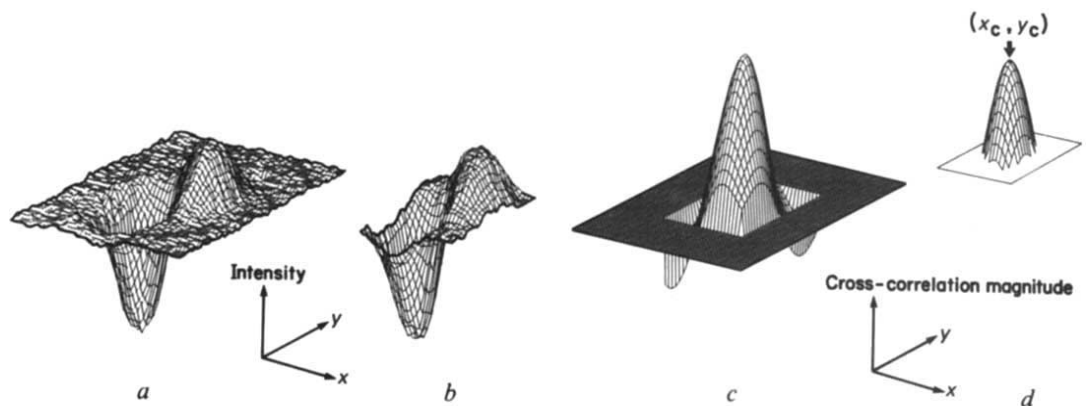
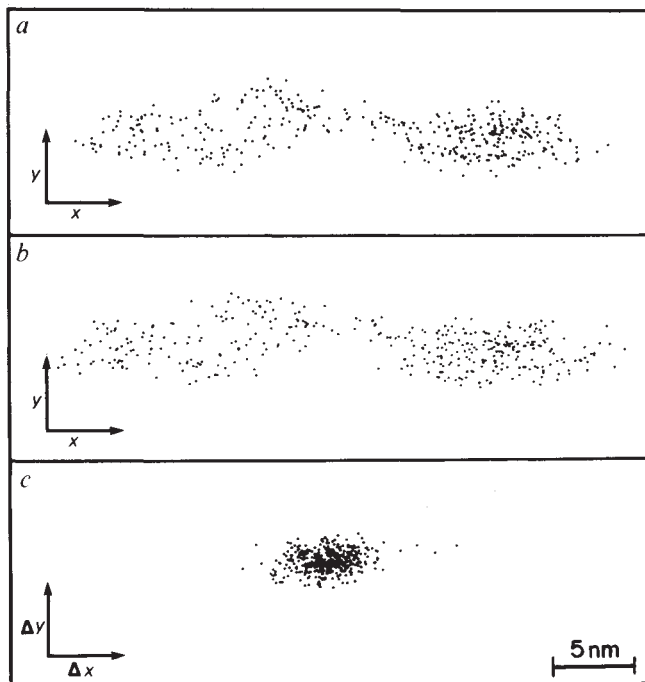


Fig. 2 The precision of the bead position measurements determined by analyses on immobilized beads. *a, b*, The positions measured for each of two beads in a sequence of 390 video frames (at 30 frames per second). For both the bead in *a* and that in *b*, the position data display a drift from the right to the left with time. *c*, The frame-by-frame difference of the bead coordinates shown in *a* and *b*. The *x* coordinate of each point is the *x* coordinate of a position measurement in *a* minus the *x* coordinate of the position measurement in *b* from the same video frame. The same is true of the *y* coordinates. Scale bar, 5 nm.

Methods. 190-nm diameter carboxylated latex beads (Polysciences Inc.) adhering to a glass coverslip were surrounded by a transparent polyacrylamide gel (10% w/v) to minimize vibrational brownian motion. The video-microscope camera output was recorded by a monochrome high-resolution optical memory disk recorder (OMDR, Panasonic TQ-2025F) at standard video-frame rates (30 frames per second). Although a frame is collected every 1/30 s, the effective temporal resolution is lower because of lag in the camera tube and because interlaced video frames partially overlap in time³⁴. The horizontal resolution of the recording is ~450 television lines per picture height. Recorded video frames were displayed by the OMDR under computer control. The video signal was processed by a 2-line time-base corrector (Fortel CCD 2h-3, 455 pixels horizontal resolution) and digitized (8-bit grey scale, 512 horizontal × 480 vertical resolution) by a video-image processing computer (Hannaway) controlled by custom software. Sixteen independently digitized images of each frame were averaged to suppress OMDR playback noise. A rectangular subregion containing the bead image to be analysed was extracted from each digitized frame and stored in the computer system for subsequent determination of the bead coordinates. A kernel subregion restricted to the bead image was also extracted from a single (arbitrarily selected) frame of the recording. The position analysis calculation was performed on a VAX 11/785 computer (Digital Equipment). For image and kernel sizes of 62 × 64 and 27 × 31 pixels, respectively, the calculation required 30 s of CPU time per frame. The threshold intensity *T* was selected empirically to minimize the noise; the same value of *T* was used to analyse each frame in a given experiment. In a typical recording, 2–5% of the frames contain recording drop-outs or video-synchronization defects which distort the bead image; these frames were excluded from subsequent analyses. Magnification was calibrated against the 0.62-μm frustule spacing in the diatom *Pleurasigma angulatum*. Pixel spacing, 54 nm horizontal × 42 nm vertical. Monte Carlo analysis of the digital centroid calculation (Fig. 1*d*) shows that the r.m.s. error introduced by digitization at this spacing is <0.5 nm.



the position of the other bead relative to it (Fig. 2*c*). The standard deviations of the positions computed in this way are 1.6 nm and 0.7 nm for the *x* and *y* coordinates, respectively, suggesting that a precision of relative bead position measurements of 1–2 nm or better can be obtained with this technique.

We examined kinesin-driven bead movements over a range of ATP concentrations (2.5 μM–1 mM). Mean bead velocities increase with increasing ATP concentration, as previously reported for kinesin-driven movements of microtubules over glass surfaces^{7,17}. A bead track from a specimen with 10 μM ATP is shown in Fig. 3*a*. The data in this figure are derived from a 16.9-s-long video recording; during this period the kinesin-coated bead moved 278 nm along a nearly vertically oriented microtubule.

Figure 3*b, c* shows the temporal relationship of the 477 individual position measurements displayed in Fig. 3*a*. Figure 3*b* shows the progress of the bead along the microtubule as measured by the projection of the bead position along the microtubule axis; Fig. 3*c* shows the small side-to-side movements perpendicular to that axis.

In Fig. 3*b* and *c*, the random scatter of the position measurements is low; it approaches that observed in the measurements on stationary beads. In contrast, the random brownian motions of unattached beads are large even when observed with the limited temporal resolution of video measurements. The root-mean-square (r.m.s.) instantaneous velocity in the perpendicular direction is 42 nm s⁻¹ for the data of Fig. 3*c*. This is >50 times slower than the expected one-dimensional r.m.s. velocity of an unattached bead of the same size observed at 30 Hz (2.3×10^3 nm s⁻¹)¹⁸. The kinesin crossbridges that attach the bead to the microtubule must therefore be sufficiently short and/or rigid strongly to constrain thermally induced bead movements.

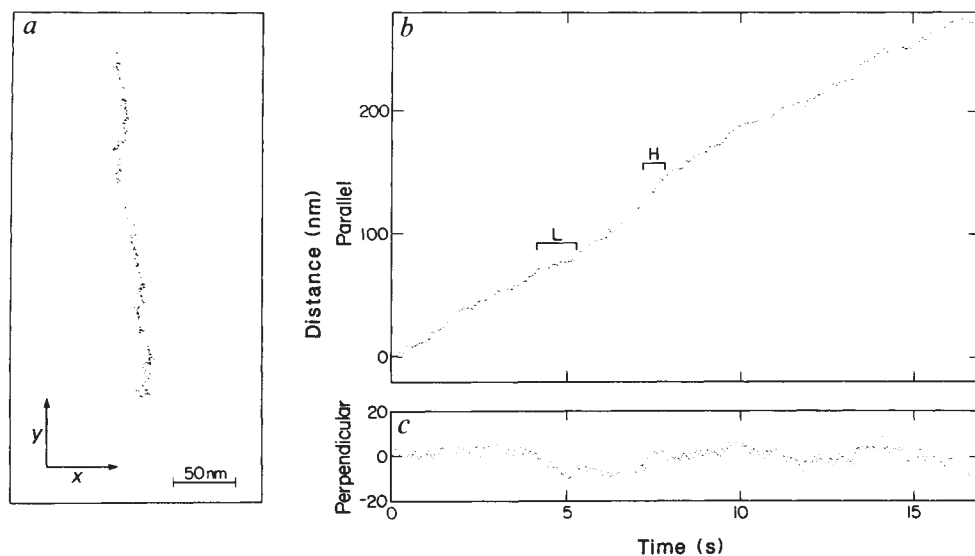
We measure the velocity of kinesin-driven bead movements along the microtubule with a much higher temporal resolution than that used previously⁴. In contrast to the uniform bead

velocities previously observed, we find that the velocity typically remains constant for an interval of only 0.1 to 2 s and then abruptly changes to a different velocity (see Fig. 3*b*, L and H), ranging from more than twice the overall mean velocity down to zero. Such variations were observed at all ATP concentrations from 10 to 1,000 μM. They could be caused by local inhomogeneities in microtubule structure, by fluctuations in the number of active force generators (compare refs 19 and 20), or by other factors. Many recordings also contained negative velocity intervals during which the bead moved backwards (usually by not more than 10 nm).

When no ATP is present, kinesin-coated beads attach to the microtubules and do not move⁴. At millimolar ATP concentrations, beads move at a maximal rate (under the assay conditions of Vale *et al.*⁹) of 500–600 nm s⁻¹. For an object moving at this velocity, the fundamental mechanical processes that produce movements (which are assumed to have a nanometre size scale) will have a frequency too high to be resolved by our instrumentation; however, we can slow such processes to an observable rate by reducing the ATP concentration.

Figure 4 shows the movement along the microtubule axis of a kinesin-coated bead in the presence of 2.5 μM ATP; its mean velocity is less than 1% of maximal. Unlike those moving at higher speeds (compare with Fig. 3), beads in this very low-velocity regime show two distinct types of motion. During some periods, movement occurs smoothly with low instantaneous velocities (see Fig. 4, region A). During other periods, however, the bead remains stationary most of the time and moves forward in sudden jumps (see Fig. 4, region B). The jump duration is usually less than two frame times (67 ms), and most jumps are ~4 nm in length. The distribution of observed data-point separation distances (not shown) confirms the tendency of the beads to stop at positions separated by ~4 nm or integer multiples of that value. This regularity might be caused by features of the kinesin mechanochemical cycle in which microtubule binding

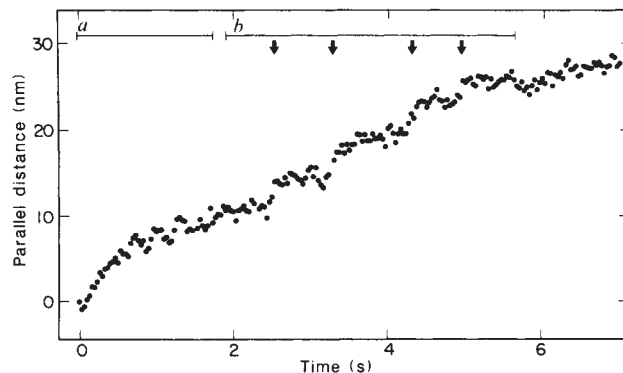
Fig. 3 Motion of a kinesin-coated bead along a microtubule in the presence of 10 μM ATP. *a*, The bead track, consisting of bead-position measurements at 30 Hz. Scale bar, 50 nm. *b*, The projection of the bead position along the direction of the bead track as a function of time. For the entire record the mean velocity is 16.1 nm s^{-1} ($\sim 3\%$ of the maximal velocity). L and H, two segments with mean velocities of 8.2 and 36.2 nm s^{-1} , respectively. *c*, The projection of the bead position along the direction perpendicular to the bead track as a function of time.



Methods. Kinesin was purified from squid optic lobes using a monoclonal antibody⁹ against the 110K polypeptide^{17,35}. We observed the motion of kinesin-coated beads on taxol-stabilized microtubules adsorbed to a coverslip silanized with dichlorooctamethyl-tetrasiloxane (Surfasil, Pierce), which made up one wall of a laminar flow chamber³⁶. Silanizing the glass minimizes the attachment of kinesin and the resultant 'gliding' of microtubules on the glass surface⁴. We added 10 μl of 2.5% (w/v) 150-nm diameter carboxylated latex beads to 90 μl kinesin in PEM buffer⁴ supplemented with 10 μM ATP, 80 mM KCl and 0.02% Triton X-100 to give a final kinesin 110K subunit concentration of 80 nM. This suspension of kinesin-coated beads was then introduced into the flow chamber and observed with a video-enhanced differential interference contrast imaging system consisting of a Zeiss ICM inverted microscope and a Dage-MTI 68 series Newvicon video camera. This system is described in detail elsewhere³⁷, with the exception of a new illuminator incorporating a fibre-optic scrambler^{34,38}. The illuminator focuses a 100-W mercury arc lamp (filtered with IR and UV absorbing filters and an interference filter with a 50-nm half width centred at 550 nm) onto a 0.75-nm diameter plastic optical fibre. The output of the fibre, which serves as a point illumination source for the microscope optics, was collected with a Zeiss infinity-corrected Neofluar objective and projected with a two-element achromat onto the condenser aperture. The focal length of the achromat was chosen to fill the aperture with an image of the fibre tip. Glan-Thompson prisms replaced the film polarizer and analyser supplied with the Zeiss DIC optics. These modifications provide significantly brighter and more even illumination, and allow higher magnification in the projection of the microscope image onto the video camera target. Video frames were recorded, digitized and processed as described in Figs 1 and 2. The *x* and *y* coordinates are the frame-by-frame differences between those of the moving bead and those of a stationary reference bead in the same frame, as in Fig. 2c. The microtubule orientation used in calculating (b) and (c) was determined by a linear least-squares fit to the points in *a*. Pixel size, 22 nm horizontal $\times$ 17 nm vertical.

Fig. 4 Motion of a kinesin-coated bead along a microtubule in the presence of 2.5 μM ATP. The plot shows the projection of the bead position along the microtubule axis as a function of time. The mean velocity along the microtubule axis over the period shown is 3.9 nm s^{-1} ($\sim 0.8\%$ of maximal velocity). Forward movement in region A occurs smoothly with a mean velocity of 5.6 nm s^{-1} . Movement in region B occurs in four discrete jumps (arrows) of length 3.7 ± 1.7 nm (mean $\pm$ s.d.).

Methods. The sample was identical to that of Fig. 3, except for a lower ATP concentration (2.5 μM). The data-analysis procedure was identical to that of Fig. 3 except that microtubule orientation was determined manually from the positions of video cursors superimposed on the centre of the microtubule image at distances of 0.5 μm on either side of the microtubule segment on which the bead motion occurred.



sites on kinesin interact specifically with a domain on periodically spaced tubulin monomers (spacing, 4 nm).

Because of the symmetry of the microtubule lattice (refs 21, 22, but see ref. 23), it is likely that identical sites for the binding of kinesin are distributed around the entire circumference of the microtubule. Despite the presence of closely spaced repeating structures over the entire microtubule surface, we do not observe any tendency for the beads to 'wander' circumferentially on the surface as they move axially over hundreds of nanometres (Fig. 3c). The absence of substantial lateral movement is intriguing when one considers the geometry of a bead attached by a kinesin crossbridge to a site on a microtubule protofilament: if the crossbridge moves from one site to a second site on an adjacent protofilament, the expected displacement of the bead centre would be between 40 and 56 nm (assuming a crossbridge length of ~ 30 nm²⁴⁻²⁶). The observed movements rarely

exceed half of these amounts (see Fig. 3c), which suggests that the bead maintains a position above a particular protofilament (or above the junction between the members of a particular protofilament pair) as it moves axially. In experiments in which several beads moved over the same segment of a microtubule, the beads followed different paths, all approximately parallel, at lateral positions over a region up to 130 nm wide. This observation may help to explain how organelles moving on microtubules pass one another without apparent collision²⁷⁻³⁰.

The simplest explanation for the restriction of lateral bead motion is that each kinesin molecule binds to sites associated with a particular protofilament (or set of adjacent protofilaments) as the bead moves along the microtubule. If a bead is attached to the microtubule with several kinesin molecules, the shift of a single molecule from one protofilament to another might produce only an insubstantial change in the circumferen-

tial position of the bead. However, both structural and energetic considerations^{4,5,11,24-26,31} suggest that the number of kinesin crossbridges linking the bead and the microtubule is small, so that even the fluctuations in the average circumferential position taken over all the crossbridges would be substantial if each crossbridge had no tendency to maintain its association with a particular protofilament.

The methods used here could be used to explore various aspects of kinesin-based movement, including characterization of the force-generating step in the mechanochemical cycle and analysis of the mechanical properties of bead attachment to microtubules in different states of the enzyme. More generally, the ability to measure molecular-scale movements of small objects in real time in the light microscope has potential for providing insights into various cellular processes.

We thank Stephen G. Turney for assistance with the instrumentation and computer software and Thomas S. Reese for his support. We also thank Wyndham Hannaway, Elliot Elson, Hong Qian and Shahid Khan for discussions. Dr Matthew Suffness (NCI) provided the taxol. Supported in part by grants from the National Institutes of Health (to M.P.S.) and a Damon Runyon-Walter Winchell Cancer Fund Fellowship (J.G.).

Received 16 October; accepted 16 December 1987.

1. Eisenberg, E. *Lect. Math. Life Sci.* **16**, 19-59 (1986).
2. Berg, H. C. & Khan, S. in *Mobility and Recognition in Cell Biology* (eds Sund, H. & Veeger, C.) 485-497 (de Gruyter, Berlin, 1983).
3. Johnson, K. A. *Rev. Biophys. biophys. Chem.* **14**, 161-188 (1985).

4. Vale, R. D., Reese, T. S. & Sheetz, M. P. *Cell* **42**, 39-50 (1985).
5. Kuznetsov, S. A. & Gelfand, V. I. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8330-8534 (1986).
6. Cohn, S. A., Ingold, A. L. & Scholey, J. M. *Nature* **328**, 160-163 (1987).
7. Porter, M. E. *et al. J. biol. Chem.* **262**, 2794-2802 (1987).
8. Khan, S., Schnapp, B. J. & Sheetz, M. P. *Biophys. J.* **49**, 415a (1986).
9. Vale, R. D. *et al. Cell* **43**, 623-632 (1985).
10. Schroer, T. A., Schnapp, B. J., Reese, T. F. & Sheetz, M. P. (in preparation).
11. Sheetz, M. P. & Spudis, J. A. *Nature* **303**, 31-35 (1983).
12. Vale, R. D., Schnapp, B. J., Reese, T. S. & Sheetz, M. P. *Cell* **40**, 559-569 (1985).
13. Vale, R. D. & Toyoshima, Y. Y. *J. Cell Biol.* **105**, 96a (1987).
14. Allen, R. D., Allen, N. S. & Travis, J. L. *Cell Motil.* **1**, 291-302 (1981).
15. Allen, R. D. *Rev. Biophys. biophys. Chem.* **14**, 265-290 (1985).
16. Howard, J. & Hudspeth, A. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3064-3068 (1987).
17. Schnapp, B. J., Chrise, B. J., Khan, S., Sheetz, M. P. & Reese, T. F. *Nature* (submitted).
18. MacDonald, D. K. C. *Noise and Fluctuations: An Introduction* (Wiley, New York, 1962).
19. Block, S. M. & Berg, H. C. *Nature* **309**, 470-472 (1984).
20. Khan, S., Meister, M. & Berg, H. C. *J. molec. Biol.* **184**, 645-656 (1985).
21. Amos, L. A. in *Microtubules* (eds Roberts, K. & Hyams, J. S.) 1-64 (Academic, London, 1979).
22. Mandelkow, E.-M. & Mandelkow, E. *J. molec. Biol.* **181**, 123-135 (1985).
23. Mandelkow, E.-M., Schultheiss, R., Rapp, R., Müller, M. & Mandelkow, E. *J. Cell Biol.* **102**, 1067-1073 (1986).
24. Miller, R. H. & Lasek, R. J. *J. Cell Biol.* **101**, 2181-2193 (1985).
25. Gilbert, S. P., Allen, R. D. & Sloboda, R. D. *Nature* **315**, 245-248 (1985).
26. Amos, L. A. *J. Cell Sci.* **87**, 105-111 (1987).
27. Koonce, M. P. & Schliwa, M. *J. Cell Biol.* **100**, 322-326 (1985).
28. Schnapp, B. J., Vale, R. D., Sheetz, M. P. & Reese, T. S. *Cell* **40**, 455-462 (1985).
29. Allen, R. D. *et al. J. Cell Biol.* **100**, 1736-1752 (1985).
30. Hayden, J. H. & Allen, R. D. *J. Cell Biol.* **99**, 1785-1793 (1984).
31. Langford, G. M., Allen, R. D. & Weiss, D. G. *Cell Motil. Cytoskel.* **7**, 20-30 (1987).
32. Ballard, D. H. & Brown, C. M. *Computer Vision* (Prentice-Hall, Englewood Cliffs, New Jersey, 1982).
33. Gonzalez, R. C. & Wintz, P. *Digital Image Processing* (Addison-Wesley, Reading, Massachusetts, 1977).
34. Inoué, S. *Video Microscopy* (Plenum, New York, 1986).
35. Steuer, E., Vale, R. D., Schnapp, B. J., Reese, T. S. & Sheetz, M. P. *J. Cell Biol.* **101**, 397a (1985).
36. Berg, H. C. & Block, S. M. *J. gen. Microbiol.* **130**, 2915-2920 (1984).
37. Schnapp, B. J. *Meth. Enzym.* **134**, 561-573 (1986).
38. Ellis, G. W. *J. Cell Biol.* **101**, 83a (1985).

Protein-disulphide isomerase and prolyl isomerase act differently and independently as catalysts of protein folding

Kurt Lang & Franz X. Schmid

Institut für Biophysik und Physikalische Biochemie,
Universität Regensburg, D-8400 Regensburg, FRG

Two enzymes are now known that catalyse slow steps in protein folding. Peptidyl-prolyl *cis-trans* isomerase¹ catalyses the *cis-trans* isomerization of Xaa-Pro peptide bonds in oligopeptides and during the refolding of several proteins^{2,3}. The other enzyme, protein-disulphide isomerase, accelerates the reactivation of reduced proteins, presumably by catalysis of thiol-disulphide exchange reactions⁴⁻⁶. Recent evidence indicates that the β -subunit of prolyl 4-hydroxylase, an enzyme involved in collagen biosynthesis, is identical with disulphide isomerase^{7,8}. On the basis of this important finding, it was suggested⁹ that disulphide isomerase accelerates protein folding, not by 'reshuffling' incorrect disulphide bonds, but in the same way as prolyl isomerase by catalysing proline isomerization which is known to be important for the folding of collagen¹⁰ and other proteins. Here we show that the catalytic activities of these two enzymes are different. Disulphide isomerase accelerates the reformation of native disulphide bonds during protein reoxidation. We find no evidence that this enzyme can catalyse the isomerization of proline peptide bonds, a reaction efficiently accelerated by prolyl isomerase. When both enzymes are present simultaneously during protein folding, they act independently of one another.

Prolyl isomerase activity is measured in a coupled assay with chymotrypsin, using the short oligopeptide *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide¹. The *p*-nitroanilide moiety can only be cleaved off rapidly by chymotrypsin when the preceding Ala-Pro peptide bond is in the *trans* conformation. At equilibrium about 10% of the peptide has a *cis* Ala-Pro bond. In the presence of a high protease concentration, hydrolysis of the *trans* peptide

occurs within a few seconds. Cleavage of the *cis* peptide, however, is limited in rate by the *cis* to *trans* isomerization of the Ala-Pro bond, which shows a time constant of 115 s at 10 °C. This reaction is efficiently catalysed by prolyl isomerase (Fig. 1a), whereas disulphide isomerase does not affect the rate of this isomerization either in the absence (Fig. 1b) or in the presence of prolyl isomerase (Fig. 1c). So under conditions where prolyl isomerase is a good catalyst, disulphide isomerase has no effect on the isomerization of the Ala-Pro bond in the test peptide.

In addition to being inactive as a prolyl isomerase when assayed with a short peptide, disulphide isomerase is unable to accelerate the slow refolding of a protein with intact disulphide bonds, such as the immunoglobulin light chain from mouse. It has been suggested that this reaction involves proline isomerization¹¹, and recently it was found to be accelerated by prolyl isomerase⁷. Figure 2a shows the catalysis of the slow refolding of this protein by prolyl isomerase. Under the same conditions, addition of disulphide isomerase at two different concentrations had no influence on the refolding of the immunoglobulin light chain (Fig. 2b). Also, disulphide isomerase did not affect the accelerated reaction when the concentration of disulphide isomerase was varied during refolding in the presence of 0.21 μ M prolyl isomerase (Fig. 2c).

Disulphide isomerase catalyses the reactivation of reduced and unfolded proteins^{4-6,12}. Figure 3 shows the acceleration of the reactivation kinetics of reduced ribonuclease A (RNase A) by disulphide isomerase. The reactivation experiments were performed in the presence of a mixture of reduced (4 mM) and oxidized glutathione (0.4 mM) at pH 7.8 and 25 °C. These are conditions where RNase A can be reactivated in the absence of disulphide isomerase, and where prolyl isomerase shows high activity towards proline-containing substrates. Addition of prolyl isomerase in variable amounts to the reoxidation mixture had no influence on the rate of reactivation of reduced RNase A (Fig. 3). In experiments where both disulphide isomerase and prolyl isomerase enzymes were present simultaneously, only a variation of the disulphide isomerase concentration affected the rate of reactivation (Fig. 3). The experiments shown in Figs 2c and 3 are complementary: they show that the two enzymes act

Fig. 1 Kinetics of *cis-trans* isomerization of the Ala-Pro peptide bond of the peptide *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide in the presence of prolyl isomerase (a), disulphide isomerase (b) and a mixture of prolyl isomerase and disulphide isomerase (c). Semi-logarithmic plots of the kinetic data are shown throughout, and the respective time constants (τ) are given below. a, Catalysis of proline isomerization in the presence of 0 nM ($\circ$; $\tau = 115$ s); 1.8 nM ($\square$; $\tau = 53$ s); 4.2 nM (Δ ; $\tau = 31$ s); 8.2 nM ($\blacksquare$; $\tau = 18$ s); and 21.2 nM ($\blacklozenge$; $\tau = 8$ s) prolyl isomerase. b, Proline isomerization in the absence ($\circ$; $\tau = 115$ s) and in the presence of 0.0027 and 0.0268 absorbance units at 280 nm of disulphide isomerase ($\bullet$). c, Isomerization in the presence of 4.2 nM prolyl isomerase (Δ ; $\tau = 31$ s) and with 0.0027 or 0.0081 absorbance units at 280 nm of disulphide isomerase added ($\blacktriangle$). All experiments were carried out in 0.1 M Tris-HCl, pH 7.8, at 10 °C.

Methods. The isomerization of the Ala-Pro bond in the test peptide *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Sigma) was measured in a coupled assay with chymotrypsin (Boehringer Mannheim); the assay based on the ability of this protease to cleave the test peptide only when the Ala-Pro bond is *trans*¹. 820 μ l chymotrypsin solution (80 μ M) in 0.1 M Tris-HCl, pH 7.8, was equilibrated at 10 °C in the spectrophotometer cell. Immediately after addition of prolyl isomerase or disulphide isomerase (in 50 μ l) the reaction was started by adding 30 μ l test peptide in dimethylsulphoxide (2.1 mM). The *trans* peptide was cleaved within the deadtime, hydrolysis of the *p*-nitrophenyl ester in the *cis* peptide is limited in rate by the *cis-trans* isomerization at Ala-Pro and was followed by the increase in absorbance at 390 nm. Prolyl isomerase was isolated from pig kidney^{1,3}; disulphide isomerase was from Genzyme (Boston, USA).

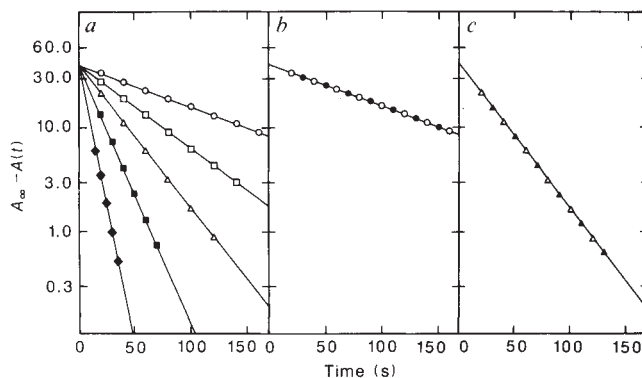


Fig. 2 Kinetics of slow refolding of the immunoglobulin light chain (disulphide bonds intact) in the presence of increasing concentrations of a, prolyl isomerase, b, disulphide isomerase and c, a mixture of both enzymes. The decrease in fluorescence at 350 nm is shown as a function of the time of refolding. The time constants (τ) derived from the semi-logarithmic plots are given below. The protein was unfolded by incubation for 1 h of light chain (40 μ M) in 5.0 M urea, 0.1 M Tris-HCl, pH 8.0, at 25 °C. Refolding of light chains (2 μ M) was achieved in 0.25 M urea and 0.1 M Tris-HCl, pH 7.8, at 10 °C. a, Variation of prolyl isomerase concentration: 0 μ M ($\circ$; $\tau = 210$ s); 0.07 μ M ($\square$; $\tau = 137$ s); 0.21 μ M (Δ ; $\tau = 90$ s); 0.34 μ M ($\blacksquare$; $\tau = 48$ s). b, Variation of disulphide isomerase concentration: ($\circ$) no disulphide isomerase, ($\bullet$) 0.0086 or 0.017 absorbance units at 280 nm. c, Refolding kinetics in the presence of 0.21 μ M prolyl isomerase alone (Δ), and with 0.0034 or 0.017 absorbance units at 280 nm of disulphide isomerase also present ($\blacktriangle$). The kinetics were measured as described³.

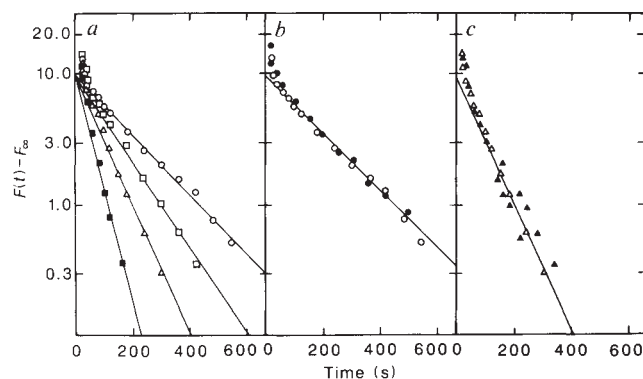
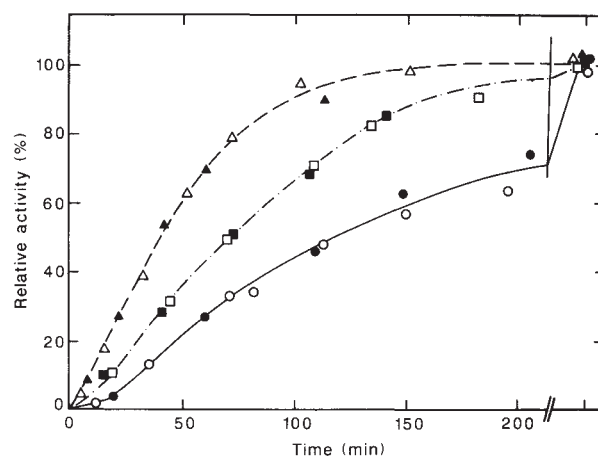


Fig. 3 Time-course of reactivation of reduced RNase A in the presence of varying concentrations of disulphide isomerase and prolyl isomerase. Reactivation was carried out in 0.1 M Tris-HCl, pH 7.8, containing 4 mM reduced and 0.4 mM oxidized glutathione and 0.1 mM EDTA at 25 °C. The final concentration of RNase A was 3.0 μ M. Reactivation in the absence of prolyl isomerase and disulphide isomerase ($\circ$); in the presence of 2 μ M prolyl isomerase ($\bullet$); and of disulphide isomerase (0.0033 absorbance units at 280 nm ($\square$); and of 0.0033 absorbance units at 280 nm disulphide isomerase and 2 μ M prolyl isomerase ($\blacksquare$); of 0.0132 absorbance units at 280 nm of disulphide isomerase (Δ); and of 0.0132 absorbance units at 280 nm of disulphide isomerase and 2 μ M prolyl isomerase ($\blacktriangle$).

Methods. Native RNase A was reduced according to Creighton¹⁴. Reactivation was initiated by dilution of 28 μ l of reduced RNase with 850 μ l 0.1 M Tris-HCl, pH 7.8, containing appropriate concentrations of reduced and oxidized glutathione and of disulphide isomerase and prolyl isomerase to give the final reactivation conditions above. A temperature of 25 °C rather than 10 °C was used in these experiments because reoxidation of reduced RNase A is extremely slow at 10 °C, but it was ascertained that disulphide isomerase is active at 10 °C as well. The activity of RNase A was assayed as described¹⁵.



specifically and independently of each other. Disulphide isomerase accelerates the reformation of correct disulphide bonds during the reactivation of reduced proteins, whereas prolyl isomerase catalyses slow steps of refolding that involve proline *cis-trans* isomerization. The failure of prolyl isomerase to affect the reactivation of reduced RNase A suggests that proline isomerization is not a rate-limiting step in the course of reoxidation of this protein.

We could not find evidence for catalytic activity of protein-disulphide isomerase in reactions that are rate-limited by proline isomerization. This holds for a small proline-containing peptide, as well as for a slow protein refolding reaction that involves proline isomerization as a rate-determining step. These reactions are accelerated by another enzyme, namely prolyl isomerase. When both disulphide isomerase and prolyl isomerase were present during refolding, no evidence for a functional inter-

dependence between the two enzymes was found. Disulphide isomerase catalyses the reactivation of reduced proteins, and prolyl isomerase can accelerate the refolding of proteins with incorrect proline isomers. Our results do not support the suggestion⁹ that disulphide isomerase increases the rate of protein folding by catalysis of proline isomerization, but they agree with existing experimental evidence and mechanistic models that assume disulphide isomerase acts by facilitating the formation of correct disulphide bonds during protein folding *in vitro* as well as *in vivo*^{5,6,13}

Received 11 November; accepted 24 December 1987.

1. Fischer, G., Bang, H. & Mech, C. *Biomed. biochim. Acta* **43**, 1101-1111 (1984).
2. Fischer, G. & Bang, H. *Biochim. biophys. Acta* **828**, 39-42 (1984).
3. Lang, K., Schmid, F. X. & Fischer, G. *Nature* **329**, 268-270 (1987).
4. Venetianer, P. & Straub, F. B. *Biochim. biophys. Acta* **67**, 166-168 (1983).
5. Freedman, R. B. *Trends biochem. Sci.* **9**, 438-441 (1984).
6. Freedman, R. B. *Nature* **329**, 196-197 (1987).
7. Koivu, J. *et al. J. biol. Chem.* **262**, 6447-6449 (1987).
8. Pihlajaniemi, T. *et al. EMBO J.* **6**, 643-649 (1987).
9. Pain, R. M. *Nature* **328**, 298 (1987).
10. Bächinger, H.-P., Bruckner, P., Timpl, R., Prockop, D. J. & Engel, J. *Eur. J. Biochem.* **106**, 619-632 (1980).
11. Goto, Y. & Hamaguchi, K. *J. molec. Biol.* **156**, 891-910 (1982).
12. Hillson, D. A., Lambert, N. & Freedman, R. B. *Meth. Enzym.* **107**, 281-294 (1984).
13. Creighton, T. E., Hillson, D. A. & Freedman, R. B. *J. molec. Biol.* **142**, 43-62 (1980).
14. Creighton, T. E. *J. molec. Biol.* **113**, 329-341 (1977).
15. Crook, E. M., Mathias, A. P. & Rabin, B. R. *Biochem. J.* **74**, 234-238 (1960).

Structural details of an adenine tract that does not cause DNA to bend

Amanda Milgram Burkhoff* & Thomas D. Tullius*†

* Department of Biology and McCollum-Pratt Institute and
† Department of Chemistry, The Johns Hopkins University,
Baltimore, Maryland 21218, USA

Runs of adenines (adenine tracts) have been implicated as the main determinant of sequence-directed DNA bending¹⁻⁴. The most widely used experimental test for bending relies on the observation that bent DNA migrates more slowly than straight DNA on a polyacrylamide electrophoresis gel¹⁻⁵. It was shown recently that the polymer (GTTTTAAAC)_n runs with normal mobility on a gel, whereas (GAAAATTTTC)_n runs more slowly and thus appears to be strongly bent⁶. The observation that these similar sequences, which differ only in the order of the adenine and thymine tracts, adopt such different shapes offers a stringent test of theories to explain DNA bending. Although the wedge model for DNA bending has recently been elaborated⁷ to explain the gel mobilities of these molecules, we wished to determine experimentally the structural basis for the difference in bending. We report here measurements of the frequency of cleavage by the hydroxyl radical at each nucleotide of cloned versions of the two polymers (see Fig. 1). We show that the TTTTAAAA sequence does not display the cleavage pattern that is associated with bent DNA⁸, whereas the AAAATTTT sequence does. The observed sequence dependence of the cleavage pattern of an adenine tract is at odds with current models^{4,7} for DNA bending, which assume that adenine tracts always adopt the same conformation.

The hydroxyl radical, which we generate by reaction of the EDTA complex of iron(II) with hydrogen peroxide, initiates cleavage of a DNA strand by abstracting a hydrogen atom from a deoxyribose in the DNA backbone⁹. The hydroxyl radical cleaves mixed-sequence DNA nearly equally at each backbone position⁹. Previously we showed⁸ that the hydroxyl-radical cleavage pattern of bent kinetoplast DNA was instead markedly sinusoidal in appearance, with the minima of the cutting pattern phased perfectly with the 3' ends of adenine tracts. The cleavage pattern of the opposite, thymine-rich, strand was similar in shape

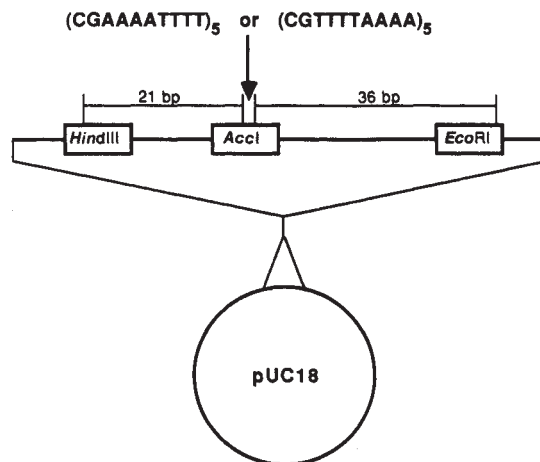


Fig. 1 Cloning oligomers of the sequences CGAAAATTTT and CGTTTTAAAA.

Methods. The partially self-complementary decanucleotides 5'-CGTTTTAAAA-3' and 5'-CGAAAATTTT-3' were synthesized by standard automated methods (Applied Biosystems) and purified by Mono-Q anion-exchange chromatography on a fast protein/peptide/polynucleotide liquid chromatography (FPLC) system. To clone an oligomer of one of the A-tract sequences, a decanucleotide duplex was ligated in the presence of *AccI*-cleaved pUC18. *Escherichia coli* (strain HB101) was transformed with the ligation mixture and plated on ampicillin plates²⁰. Clones containing plasmid with five repeats of a decamer were selected.

but shifted by one to three nucleotides in the 3' direction, indicating that the cleavage rate is correlated for nucleotides directly across the minor groove from one another. Our interpretation of the kinetoplast DNA cutting pattern therefore implicated a progressive narrowing of the minor groove from 5' to 3' along an adenine tract as a structural feature linked with DNA bending⁸. Oligo(dA)-oligo(dT) tracts have been shown by recent single crystal X-ray diffraction experiments^{10,11} to have a more narrow minor groove than mixed-sequence DNA.

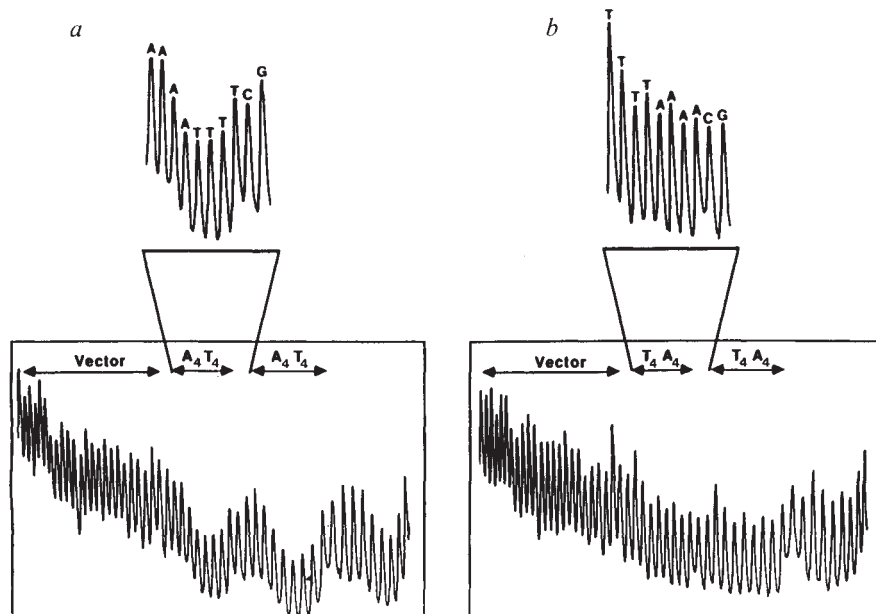
In Fig. 2 we compare the hydroxyl-radical cutting patterns of oligomers of the sequences CGAAAATTTT and CGTTTTAAAA. An important feature of the design of the experiment is that the restriction fragments studied contain mixed-sequence DNA from the cloning vector flanking the adenine tract oligomers. This internal control allows us to compare the cutting frequencies of an adenine tract and 'normal' DNA. We find that each repeat of a particular adenine-tract oligomer has the same cutting pattern, illustrating the reproducibility and sensitivity of the method.

The CGAAAATTTT sequence shows a clear sinusoidal cutting pattern, whereas the adjacent vector DNA is cut with a fairly even frequency that is higher than the minima in the CGAAAATTTT tracts (Fig. 2a). In contrast, the CGTTTTAAAA sequence (Fig. 2b) exhibits an almost even cutting pattern, similar in intensity to that of the flanking mixed-sequence DNA. We conclude that the adenine tracts in (CGTTTTAAAA)_n adopt a different conformation from the adenine tracts in (CGAAAATTTT)_n.

What are the different structures of the two sequences? Our interpretation of the cleavage patterns is that the minor groove of the AAAATTTT sequence periodically opens and closes in width, like bent kinetoplast DNA⁸, and that the TTTTAAAA sequence has a minor groove as wide as in mixed-sequence DNA. Several lines of evidence support these conclusions. The crystal structure of the dodecamer CGCGAATTCGCG shows that the minor groove in the AATT sequence is narrower than in the CG regions¹². Two of the structural characteristics associ-

Fig. 2 Densitometer scans of autoradiographs showing the hydroxyl-radical cleavage patterns of the T_4A_4 and A_4T_4 oligomers. Arrows above the densitometer scans indicate oligomer and vector sequences. Two of the five repeats of each adenine-tract sequence are shown. A scan at expanded scale of one repeat of each sequence is shown at the top, with the sequence marked. *a*, The cleavage pattern of the A_4T_4 oligomer is sinusoidal, reminiscent of the cleavage pattern of kinetoplast DNA⁸. The cutting frequency decreases from the 5' adenine to the 5' thymine, and then increases. The cleavage pattern on the other strand (not shown) is nearly identical in shape but offset by 2–3 nucleotides to the 3' direction⁸, indicating that the observed changes in cleavage frequency are associated with structural features of the minor groove¹⁸. *b*, The T_4A_4 oligomer has a relatively even cleavage pattern, not much different from that of flanking vector DNA. The only exception is the first T in each oligomer, which is cut at higher frequency than surrounding sequences, perhaps indicating an unusual structural environment for this nucleotide.

Methods. Purified plasmids²¹ were linearized with *Hind*III and *Eco*RI and 3'-labelled²⁰ with ³²P at the *Hind*III end. After reaction with hydroxyl radical⁸, the reaction products were denatured and electrophoresed on an 8% polyacrylamide/8 M urea gel²⁰. Bands were assigned by reference to products of a Maxam-Gilbert G + A sequencing reaction²² run on an adjacent lane of the gel. Hydroxyl-radical cleavage leaves 5' phosphate ends, the same as Maxam-Gilbert chemistry²². The assigned base corresponds to the deoxyribose that was attacked by the hydroxyl radical.



ated with the narrow minor groove in the AATT sequence are the large positive propeller twists of the base pairs¹² and an ordered water structure ('spine of hydration')¹³ in the minor groove. Changes in sequence that affect these structural features lead to widening of the minor groove¹². For example, guanine, with its 2-amino group in the minor groove, interferes with the spine of hydration^{12,13}. Interruption of an adenine tract by a guanine (AAGAA) abolishes bending⁴.

In the case of the two adenine tracts considered in this paper, the sequence differences are more subtle. One obvious difference is the presence of a 5'-TpA-3' step in the sequence with normal gel mobility⁶, instead of a 5'-ApT-3' step in the bent sequence. Recent calculations^{14,15}, supported by the Calladine-Dickerson rules^{16,17} for B-DNA structure, have indicated that the TpA step will have a lower propeller twist because of steric clash in the minor groove between cross-strand adenines. A smaller propeller twist would lead to a wider minor groove¹². Nuclease digestion studies¹⁸ suggest that the TpA step has a wider minor groove than the ApA or ApT step. It was shown that interruption of an adenine tract by a single thymine (the sequence AATAA, with one TpA step) eliminates bending⁴.

Our results call into question key assumptions of the newest version of the wedge model⁷ that was designed to account for the gel mobility differences⁶ of $(GAAATTTC)_n$ and $(GTTTAAAAC)_n$. The original wedge model¹⁹ postulated that the ApA dinucleotide forms a 'wedge' which, when phased with the helix screw, causes DNA to bend. By assuming particular values of roll and tilt angles for an AA dinucleotide wedge, Ulanovsky and Trifonov⁷ found that simple vector addition of the roll and tilt of individual AA dinucleotides sufficed to rationalize the gel mobilities of a variety of adenine-tract sequences. If these vectors happen to cancel (as was proposed⁷ for $(GTTTAAAAC)_n$), the DNA molecule is predicted to lack global curvature. We emphasize, though, that according to the wedge model^{7,19} the local conformation of an AA dinucleotide is the same for any sequence, curved or straight; it is only the summation of these local structural anomalies that yields curvature.

In contrast, the gradient of hydroxyl-radical cutting frequency along an adenine tract in bent DNA (Fig. 2*a*, and ref. 8) is

evidence that each adenine is in a slightly different structural environment. Also we have shown here that an adenine tract is capable of adopting a conformation different from that associated with bent DNA⁸. Therefore, no cancellation of wedge angles⁷ is necessary to explain the normal gel mobility of the CGTTTAAAA polymer. Our experiments show that this sequence lacks a particular structural feature characteristic of bent DNA, and for this reason is unlikely to be bent.

The details of the structures that are actually adopted by the two sequence isomers are not as important as the observation that there is a clear dependence of the structure of an adenine tract on its sequence context. We offer one possible explanation for this structural difference, a change in groove shape, which is supported by other structural data.

The roll and tilt angles (8.4° and 2.4°) of the most recent version of the wedge model⁷ were calculated by assuming that the wedge vectors of GTTTTAAAAC cancel. As we have shown here that the adenine tracts in CGTTTAAAA do not adopt the same structure as other adenine tracts, quantitative application of these wedge angles⁷ to other sequences should be reassessed. Detailed structural studies of a wide variety of DNA sequences will be necessary before we can construct a general theory to describe DNA bending.

We thank Daniel Burkhoﬀ, Victor Corces and Paul Englund for their guidance and valuable discussions, and David Mold for synthesizing the oligonucleotides. This research was supported by grants from the Searle Scholars Program of the Chicago Community Trust, the MacArthur Foundation, and the NCI and the BRSG Program of the NIH.

Received 8 September; accepted 4 January 1988.

1. Wu, H.-M. & Crothers, D. M. *Nature* **308**, 509–513 (1984).
2. Hagerman, P. J. *Biochemistry* **24**, 7033–7037 (1985).
3. Dickmann, S. & Wang, J. *J. molec. Biol.* **186**, 1–11 (1985).
4. Koo, H.-S., Wu, H.-M. & Crothers, D. M. *Nature* **320**, 501–506 (1986).
5. Marini, J. C., Levene, S. D., Crothers, D. M. & Englund, P. T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7664–7668 (1982).
6. Hagerman, P. J. *Nature* **321**, 449–450 (1986).
7. Ulanovsky, L. E. & Trifonov, E. *Nature* **326**, 720–722 (1987).
8. Burkhoﬀ, A. M. & Tullius, T. D. *Cell* **48**, 935–943 (1987).
9. Tullius, T. D. & Dombroski, B. A. *Science* **230**, 679–681 (1985).

10. Nelson, H. C. M., Finch, J. T., Luisi, B. F. & Klug, A. *Nature* **330**, 221–226 (1987).
11. Coll, M., Frederick, C. A., Wang, A. H.-J. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8385–8389 (1987).
12. Fratini, A. V., Kopka, M. L., Drew, H. R. & Dickerson, R. E. *J. biol. Chem.* **257**, 14686–14707 (1982).
13. Drew, H. R. & Dickerson, R. E. *J. molec. Biol.* **151**, 535–556 (1981).
14. Chuprina, V. P. *FEBS Lett.* **186**, 98–102 (1985).
15. Chuprina, V. P. *Nucleic Acids Res.* **15**, 293–311 (1987).
16. Calladine, C. R. *J. molec. Biol.* **161**, 343–352 (1982).
17. Dickerson, R. E. *J. molec. Biol.* **166**, 419–441 (1983).
18. Drew, H. R. & Travers, A. *Cell* **37**, 491–502 (1984).
19. Trifonov, E. N. & Sussman, J. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3816–3820 (1980).
20. Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
21. Tullius, T. D., Dombroski, B. A., Churchill, M. E. A. & Kam, L. *Meth. Enzym.* **155**, 537–558 (1987).
22. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).

Multiple liquid crystal phases of DNA at high concentrations

Teresa E. Strzelecka, Michael W. Davidson & Randolph L. Rill

Department of Chemistry and Institute of Molecular Biophysics,
The Florida State University, Tallahassee, Florida 32306-3006, USA

DNA packaging *in vivo* is very tight, with volume concentrations approaching 70% w/v in sperm heads, virus capsids and bacterial nucleoids^{1–3}. The packaging mechanisms adopted may be related to the natural tendency of semi-rigid polymers to form liquid crystalline phases in concentrated solutions^{4–8}. We find that DNA forms at least three distinct liquid crystalline phases at concentrations comparable to those *in vivo*, with phase transitions occurring over relatively narrow ranges of DNA concentration. A weakly birefringent, dynamic, 'precholesteric' mesophase with microscopic textures intermediate between those of a nematic and a true cholesteric phase forms at the lowest concentrations required for phase separation. At slightly higher DNA concentrations, a second mesophase forms which is a strongly birefringent, well-ordered cholesteric phase with a concentration-dependent pitch varying from 2 to 10 μm . At the highest DNA concentrations, a phase forms which is two-dimensionally ordered and resembles smectic phases of thermotropic liquid crystals observed with small molecules.

Although specialized proteins are involved in many DNA packaging processes, descriptions of mechanisms of DNA packaging must also consider the intrinsic tendency of the stiff DNA chain to fold when confined to a small volume. Stiff non-electrolyte polymers form highly ordered, liquid crystalline phases above a critical concentration dependent on the persistence length^{4–8}. Formation of nematic or cholesteric liquid crystalline phases by semi-rigid polymers, such as polybenzyl-L-glutamate, in organic solvents has been studied in detail^{7,8}, and was predicted theoretically by Onsager⁴, Flory^{5,6} and others. Much less is known about the behaviour of semi-rigid polyelectrolytes. Because DNA, fibrous proteins and certain polysaccharides are polyelectrolytes involved in numerous supra-molecular associations *in vivo*, an understanding of the phase behaviour of these biopolymers is of fundamental biological importance.

Ordering of semi-rigid polymers at high concentrations occurs spontaneously to minimize the macromolecular excluded volume^{4–8}. Semi-rigid polyelectrolyte behaviour is complicated by charge-shielding requirements. A strong polyelectrolyte is surrounded by a counterion layer which determines the effective axial ratio and excluded volume^{9–13}. Because polymer phase behaviour depends on the effective polymer dimensions, the critical concentration for DNA ordering is a sensitive function of ionic strength and counterion type.

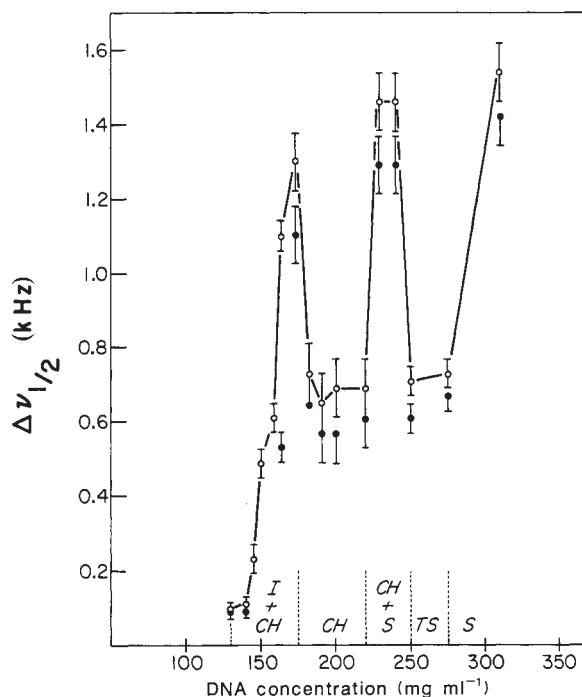


Fig. 1 Dependence of solid state ^{31}P NMR linewidth on DNA concentration suggesting multiple phase transitions. DNA fragments of average length 146 base pairs ($\sim 500 \text{ \AA}$) and with a narrow distribution about this length were prepared by digestion with micrococcal nuclease of calf-thymus chromatin previously depleted of histone H1, and by subsequent deproteinization¹⁶. Short DNA molecules are studied as a preliminary step to understanding the natural behaviour of DNA at *in vivo* concentrations. Such defined samples are useful because effects of relative molecular mass heterogeneity are minimized and phase transitions are sharp and kinetically rapid. Using DNA fragments of this size is appropriate because the driving forces are the same for ordering of high axial ratio, rod-like molecules and long semi-flexible polymers of identical composition^{5,6}. Open circles: data taken at 30°C ; filled circles: data taken at 50°C . Spectra were obtained on non-spinning samples at a phosphorus frequency of 61.3 MHz on the 'Seminole', an in-house constructed multinuclear Fourier transform (FT) spectrometer equipped with wide-bore superconducting solenoid and quadrature detection, modified for solid-state applications. Sweep width was $\pm 25 \text{ kHz}$, pulse repetition times were 3 s, and 600 scans were added for each spectrum. Gated proton decoupling was applied during data acquisition.

Previous studies demonstrated that aqueous solutions of persistence length DNA ($\sim 500 \text{ \AA}$), with 0.3 M NaCl as the supporting electrolyte, form biphasic liquid crystalline solutions containing spherulites at DNA concentrations (C_D) of 130–170 mg DNA ml^{-1} at room temperature, and become fully liquid crystalline at higher C_D (refs 14–16). Phase transition boundaries were in good agreement with predicted rigid rod behaviour when DNA was treated as a scaled rod with an effective radius of $21.5 \text{ \AA}$ at this ionic strength¹⁶. Here we report the phase behaviour of DNA solutions over a range of C_D s from 100 to 350 mg DNA ml^{-1} , as determined by solid-state ^{31}P NMR spectroscopy, polarized light microscopy and by electron microscopy.

Solutions showed a uniform liquid crystalline phase at C_D s from 170 to 220 mg ml^{-1} . A C_D of 170 mg ml^{-1} corresponds to an effective DNA volume fraction of 0.76, assuming $21.5 \text{ \AA}$ for the effective DNA radius¹⁶. Extrapolation to higher C_D s yields an apparent effective volume fraction > 1.0 for C_D s exceeding $\sim 230 \text{ mg ml}^{-1}$. As DNA solutions were prepared with C_D s over

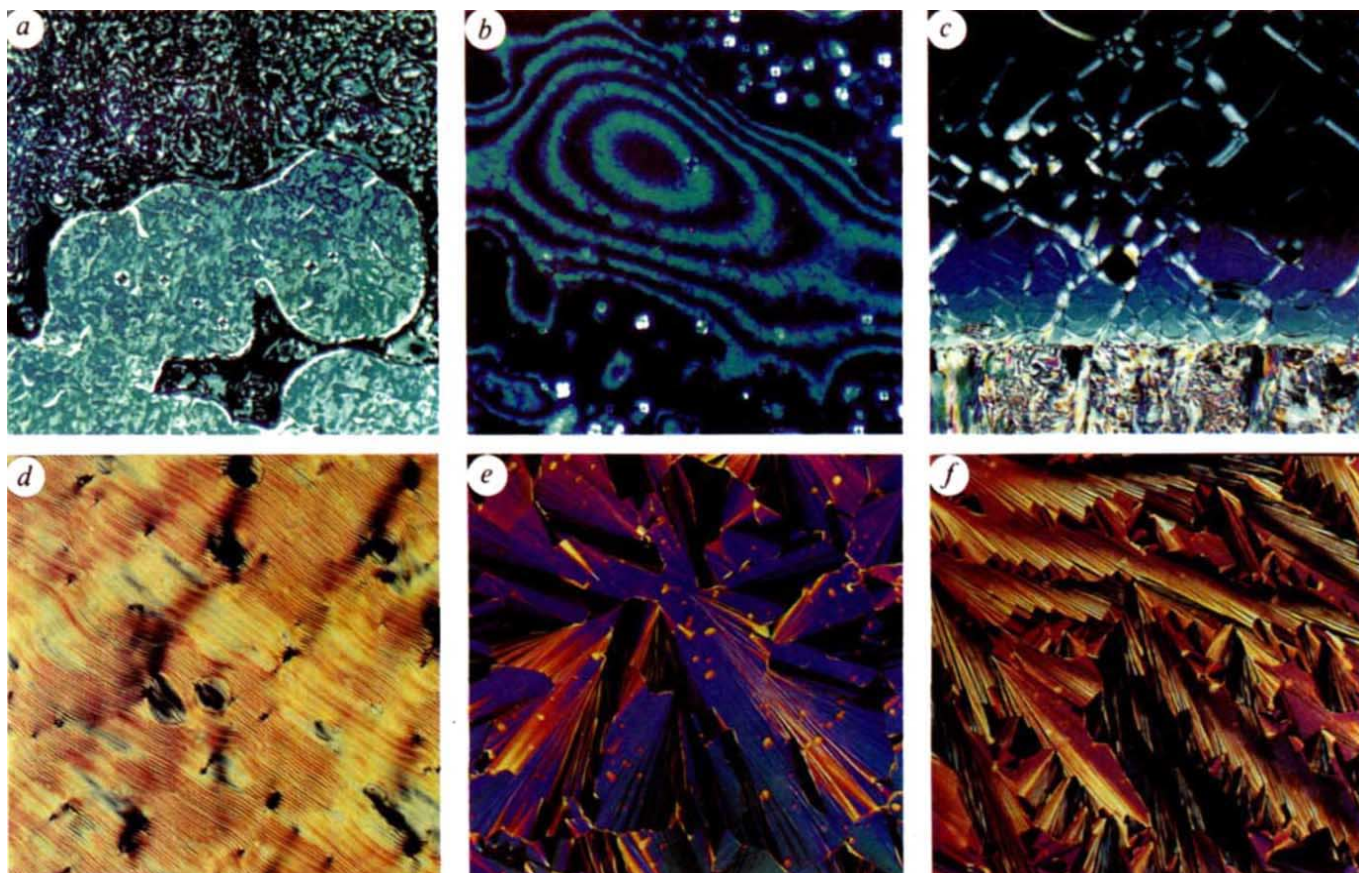


Fig. 2 Multiple liquid crystalline DNA phases observed by polarized light microscopy. A sample of 100 mg ml^{-1} DNA in 0.25 M ammonium acetate was placed on a slide under a partially sealed coverslip so that slow evaporation created a continuous concentration gradient. Samples were observed through crossed polarizers with a Nikon Optiphot-Pol microscope under $3,200 \text{ K}$ tungsten-halide illumination and photographed on Fujichrome 64-T professional film. *a, c*, Low magnification views illustrating sharp transition zones between precholesteric and cholesteric domains (*a*) and between cholesteric and smectic-like domains (*c*). Magnification $\sim \times 40$. *b, d-f*, Higher magnification views of the three major phases. *b*, Weakly birefringent, diffuse ring texture of 'precholesteric' phase. *d*, Highly birefringent, fringe or chevron texture of magnetically aligned cholesteric phase. *e*, Focal-conic-fan texture of smectic-like phase. *f*, 'Pleated ribbon' texture of most concentrated smectic-like phase near open end of coverslip. Magnifications $\sim \times 100$ in *b, e, f*; $\times 200$ in *d*.

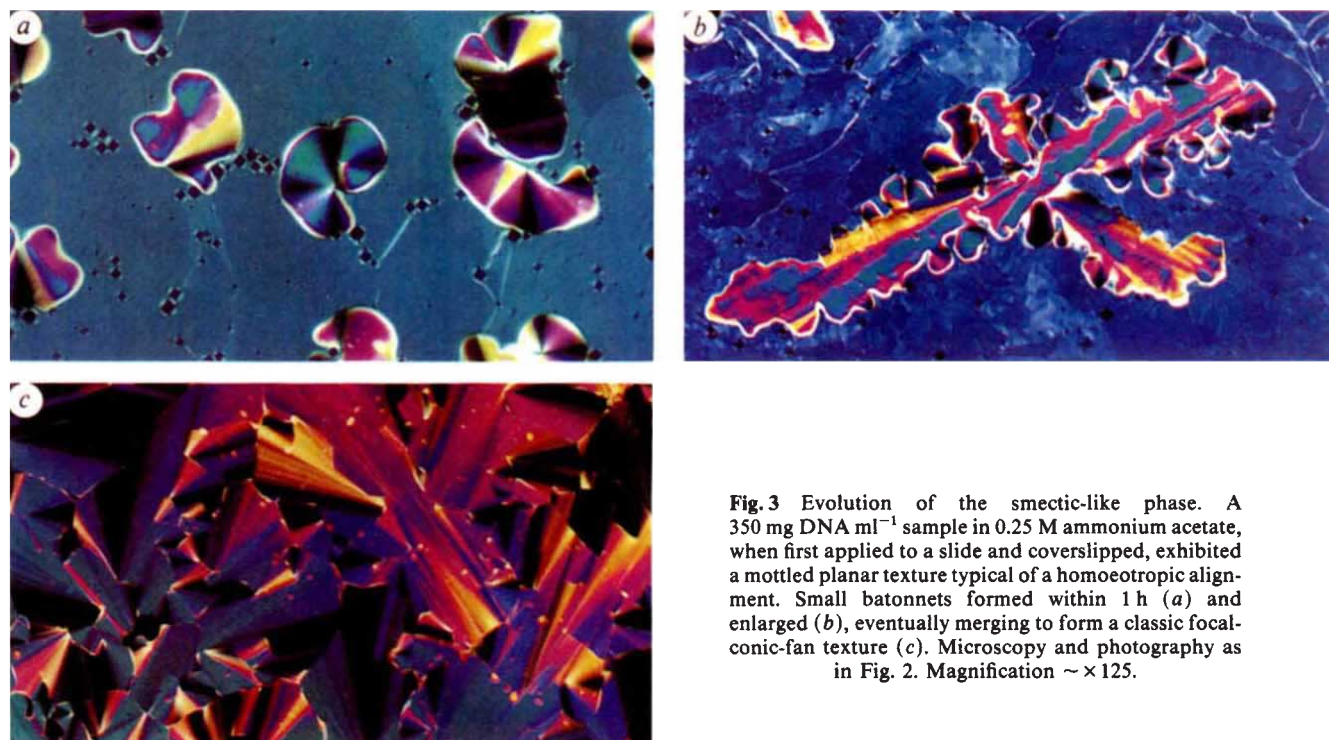


Fig. 3 Evolution of the smectic-like phase. A $350 \text{ mg DNA ml}^{-1}$ sample in 0.25 M ammonium acetate, when first applied to a slide and coverslipped, exhibited a mottled planar texture typical of a homoeotropic alignment. Small batonnets formed within 1 h (*a*) and enlarged (*b*), eventually merging to form a classic focal-conic-fan texture (*c*). Microscopy and photography as in Fig. 2. Magnification $\sim \times 125$.

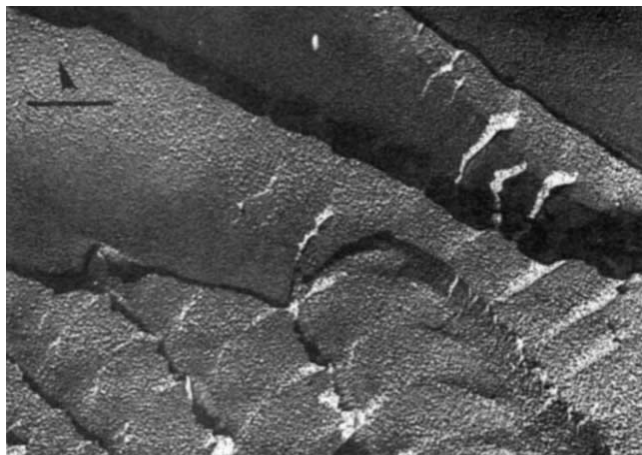


Fig. 4 Freeze-fracture-etch electron microscopy illustrating the layered structure of the smectic-like, high-concentration DNA phase. Low magnification view shows cleavage 'steps' (scale bar, 1 μm). The direction of shadowing is indicated by the arrow. DNA molecules appear to lie 'end-on' to the viewer and approximately perpendicular to the layer planes, as expected for a smectic phase. The 'end-on' view is readily distinguished from 'side-on' views obtained on examination of cholesteric phases, which tend to shear across cholesteric planes (M. W. Davidson and R. L. Rill, unpublished observations). Samples were placed on gold grids, quick-frozen between gold planchets using propane jets, fractured at -120°C , etched at -100°C for 2 min, and platinum-carbon shadowed at 45°C (Balzers BAE 360). Micrographs of replicas were obtained on a Jeol 100CX in transmission mode at 80 kV acceleration voltage.

300 mg ml^{-1} , there must be mechanisms for reducing the effective DNA volume fraction.

Three potential mechanisms for reducing the effective DNA volume are evident: (1) a change to a more ordered phase, (2) a contraction of the counterion layer, or (3) a change in DNA conformation/hydration. A phase change above C_{DS} of 220 mg ml^{-1} was indicated by a change in ^{31}P resonance linewidth and shape (Fig. 1). At lower C_{DS} , in the biphasic region, the resonance linewidth increased progressively with C_{D} until the solution became fully liquid-crystalline, then dramatically sharpened. Concurrently, the lineshape changed from lorentzian to an asymmetric form expected from an aligned sample with some rotational averaging. Magnetic alignment of liquid crystalline DNA phases with the long molecular axes perpendicular to the field was reported previously^{16,17}. Magnetic alignment of the mesophase in a biphasic solution is prevented because surface tension restricts the mesophase to spherulites in which DNA molecules are oriented tangentially to the spherical surface¹⁸. By analogy, we attribute lineshape changes at C_{DS} slightly above 220 mg ml^{-1} to formation of a biphasic solution in which two liquid-crystalline phases coexist. Magnetic alignment may be ineffective because of dynamic exchange of molecules between phases with different alignment properties. This conclusion was enforced by the behaviour at yet higher C_{DS} (Fig. 1). The resonance remained broad at C_{DS} from 229 to 240 mg ml^{-1} , then narrowed again at C_{DS} of 250 and 275 mg ml^{-1} , suggesting formation of a second uniform, aligned phase. Resonance broadening indicative of another phase transition was again observed at a C_{D} of 309 mg ml^{-1} .

Interpretations of lineshape changes in terms of phase transitions were supported by microscope examination. Liquid-crystalline phases are traditionally characterized by their microscopic textures as observed through crossed polarizers¹⁹⁻²¹. Cholesteric liquid crystals, in which the long axes of the molecules lie in pseudo-planes that are slightly twisted with respect to each other, have periodic variations in refractive index and fringe patterns with spacings of $P/2$, where P is the cholesteric pitch. Many smectic liquid crystals show focal-conic fan textures due to ordering in a second dimension, and can be easily distinguished from nematic or cholesteric forms^{20,21}.

Three distinct phases were observed simultaneously by placing a $100\text{ mg DNA ml}^{-1}$ sample on a slide and partially sealing the coverslip, allowing slow evaporation to create a gradient in C_{D} across the slide (Fig. 2a,c). The low C_{D} mesophase was weakly birefringent and gave roughly circular patterns of broad alternating light and dark lines when viewed through crossed polarizers (Fig. 2a,b). Rapid, small-scale changes in birefringence due to fluctuations in the nematic director were noted at high magnification ($\times 400$). The weak birefringence and director fluctuations imply a dynamic charac-

ter of this phase. Thicker specimens of this phase had an 'oily streak' texture and could be magnetically aligned, but were easily perturbed when removed from the field.

We tentatively refer to the first mesophase as a 'precholesteric' phase. A weakly birefringent precholesteric mesophase in moderately concentrated samples of high relative molecular mass DNA was recently reported¹⁹. The textures of the latter phase were attributed to a three-dimensional helicoidal arrangement of DNA molecules somewhat analogous to the molecular arrangement in blue phases of small molecules. Microscopic textures we observed for the precholesteric phase of short DNA were unlike those reported¹⁹, suggesting that evolution of DNA phases may be molecular-length dependent.

The periodic fringe and 'oily streak' textures of the first mesophase are reminiscent of a cholesteric phase, but they differed significantly from the classic cholesteric textures of the intermediate DNA mesophase, which was highly birefringent with very regular, finely spaced fringe patterns and was extensively aligned in a magnetic field (Fig. 2d). Freeze-fracture-etch electron microscopy and optical diffraction studies have confirmed that this intermediate phase is, in fact, cholesteric (M. W. Davidson and R. L. Rill, manuscript in preparation).

Smectic phases have two-dimensional order, with the molecules arranged in planes at a defined angle to the preferred orientation direction of the long molecular axes. There are two common 'natural' textures of smectic A and several other smectic phases—a uniform, homeotropic texture with molecules oriented nearly perpendicular to the slide surface, and the focal-conic fan texture^{20,21}. Focal-conic and striated fan textures were observed for the highest concentration mesophase(s) in controlled drying experiments (Fig. 2c,e,f). Close examination of a $350\text{ mg DNA ml}^{-1}$ sample confirmed that DNA forms a mesophase with molecular layering resembling that found in smectic phases of small molecules. Specimens observed immediately after placement on a slide usually exhibited a nearly uniform homeotropic texture. Batonnets initiated after a few hours, enlarged, and eventually merged to yield a classic focal-conic fan texture (Fig. 3). Examination in the electron microscope of freeze-fracture-etch replicas confirmed that DNA molecules were arranged in layers, and that the long molecular axes were approximately perpendicular to the layer planes (Fig. 4). More definitive data are required to unambiguously relate this DNA phase to conventional smectics. Using light and electron microscopy we observed several other textures in samples of $\sim 300\text{--}400\text{ mg ml}^{-1}$ (for example, compare Fig. 2e,f), suggesting the existence of other higher-order phases. As DNA is chiral and packs hexagonally in crystals²², chiral phases could form which are analogous to the chiral smectic C phase, or to more ordered chiral hexatic phases of small molecules (for example, chiral smectic G or I (refs 20, 21). These observations demon-

strate that DNA *in vitro* can assume multiple-packing arrangements which can be readily altered over small ranges of concentration. We expect that the phase behaviour of DNA *in vitro* will also be a sensitive function of the ionic environment. Significant modulation of DNA packing *in vivo* may be accomplished by small changes in DNA concentration or counterion atmosphere. Similar principles may apply to the ordering of other biological polyelectrolytes.

We thank Dr Henry Aldrich of the University of Florida for assistance with freeze-fracture-etch methods and Drs David Van Winkle and Francoise Livolant for advice on liquid crystal theory. This work was supported in part by the NIH.

Received 23 November; accepted 16 December 1987.

1. Du Praw, E. J. *DNA and Chromosomes* (Holt, Rinehart & Winston, New York, 1970).
2. Earnshaw, W. C. & Casjens, S. R. *Cell* **21**, 319-331 (1980).
3. Sipski, W. & Wagner, T. E. *Biopolymers* **16**, 573-582 (1977).
4. Onsager, L. *Ann. N.Y. Acad. Sci.* **51**, 627-659 (1949).
5. Flory, P. J. *Proc. R. Soc. A* **234**, 60-89 (1956).
6. Flory, P. J. in *Polymer Liquid Crystals* (eds. Ciferri, A., Krigbaum, W. R. & Meyer, R. B.) 103-112 (Academic, New York, 1982).
7. Robinson, C. *Trans. Faraday Soc.* **52**, 571-592 (1956).
8. Miller, W. A. *Rev. Phys. Chem.* **29**, 519-535 (1979).
9. Manning, G. S. Q. *Rev. Biophys.* **11**, 170-246 (1978).
10. Stigter, D. *Biopolymers* **16**, 1435-1448 (1977).
11. Stigter, D. *J. Phys. Chem.* **82**, 1603-1606 (1978).
12. Brian, A. A., Frisch, H. L. & Lerman, L. S. *Biopolymers* **20**, 1305-1328 (1981).
13. LeBret, M. & Zimm, B. H. *Biopolymers* **23**, 271-285 & 286-312 (1984).
14. Rill, R. L., Hilliard, P. R. & Levy, G. C. *J. Biol. Chem.* **258**, 250-256 (1983).
15. Rill, R. L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 342-346 (1986).
16. Strzelecka, T. E. & Rill, R. L. *J. Am. Chem. Soc.* **109**, 4513-4518 (1987).
17. Brandes, R. & Kearns, E. R. *Biochemistry* **25**, 5890-5895 (1986).
18. Lerman, L. S. *Cold Spring Harbor Symp. quant. Biol.* **38**, 59-74 (1974).
19. Livolant, F. *J. Physique* **48**, 1051-1066 (1987).
20. Demus, D. & Richter, L. *Textures of Liquid Crystals* (Verlag Chemie, New York, 1978).
21. Gray, G. W. & Goodby, J. W. *G. Smectic Liquid Crystals* (Heyden, Philadelphia, 1984).
22. Lerman, L. S., Wilkerson, L. S., Venable, J. H. & Robinson, B. H. *J. molec. Biol.* **108**, 271-293 (1976).

GUIDE TO AUTHORS

Please follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers for whom English is a second language and those who work in other fields. Please write clearly and simply, avoiding unnecessary technical terminology. *Nature's* staff will edit manuscripts to those ends if necessary. Contributors should check their proofs carefully.

Because of the competition for space, many of the papers submitted for publication cannot be accepted. For this reason, and because brevity is a great assistance to readers, papers should be as brief as is consistent with intelligibility. Please note that one printed page of *Nature*, without diagrams or other interruptions of the text, has fewer than 1,300 words.

CATEGORIES OF PAPER

Review Articles survey recent developments in a field. Most are commissioned, but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance but must not exceed six pages of *Nature*.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text or more than six display items (figures and tables) and should not occupy more than five pages of *Nature*.

Articles should be accompanied by a heading of 50-80 words written to advertise their contents in general terms, to which editors will pay particular attention. A heading (printed in *italics*) is not an abstract and should not usually contain numbers or measurements. The study should be introduced in more detail in the first two or three paragraphs, which should also briefly summarize its results and implications.

Articles may contain a few subheadings of two or three words. The meaning of the text should not depend on the subheadings, whose function is to break up the text and to point to what follows. There should be fewer than 50 references.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should not have more than 1,000 words of text or more than four display items and should not occupy more than two pages of *Nature*. The first paragraph should describe, in not more than 150 words, the origin and chief conclusions of the study. Letters should not have subheadings or more than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned, most are unsolicited. They are normally between one and four pages of *Nature* in length.

News and Views articles are intended to inform non-specialist readers about a recently published advance. Suggestions should be made to the News and Views Editor. Illustrations are welcome. Proposals for meeting reports should be agreed in advance.

Scientific Correspondence is for the discussion of scientific matters, including contributions published in *Nature*. Priority is given to contributions of less than 500 words and five references. Figures are welcome.

Submission. All manuscripts may be submitted to either London or Washington. They should not be addressed to editors by name. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax (VAT).

GENERAL

All contributions submitted for publication in *Nature* should conform with these rules.

Manuscripts should be typed, double-spaced, with a good-quality printer, on one side of the paper only. Four copies are required, each accompanied by lettered artwork. Four copies of half-tones should be included. Reference lists, figure legends and so on should be on separate sheets, all of which should be double-spaced and numbered. Copies of relevant manuscripts in press or submitted for publication elsewhere should be included, clearly marked as such. Revised and resubmitted manuscripts should also be clearly marked as such and labelled with their reference numbers.

Titles should say what the paper is about with the minimum of technical terminology and in fewer than 80 characters. Authors should avoid active verbs, numeric values, abbreviations and punctuation and should include one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name, figure number and, when known, manuscript reference number. Original artwork should be unlettered. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are permitted only if the parts are closely related, either experimentally or logically. Suggestions for cover illustrations, with captions, are welcome. Original artwork (and one copy of the manuscript) will be returned when a manuscript cannot be published.

Colour Artwork. A charge of £500 a page is made for 4-colour figures. Inability to meet these costs will not prevent the publication of essential colour figures if the circumstances are explained. **Figure legends** must not exceed 300 words and ideally should be much shorter, and should use telegraphic form wherever possible. The figure should be described first, then, briefly, the method. Reference to a method described elsewhere is preferable to a full description. Methods should not be described in detail in the text.

References should be numbered sequentially as they appear in the text, followed by those in tables and finally those in the figure legends. Reference numbers apply only to papers published or in the press, and a different number should be given to each paper cited. All other forms of reference (including unrefereed abstracts) should be included in the text as personal communication, manuscript in preparation or preprint (with number and institution where appropriate). Text should not be included in the reference list. References should be abbreviated according to the *World List of Scientific Periodicals* (fourth edition, Butterworth, London, 1963-65). The first and last page numbers should be cited. Reference to books should clearly indicate the publisher and the date and place of publication.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided as much as possible and, when used, defined. Editors will shorten words if necessary. In case of doubt, SI units should be used.

Footnotes should not be used except for changed addresses or to identify the corresponding author if different from the first-named.

Acknowledgements should be brief. Grant numbers and contribution numbers are not allowed.

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnType ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (____) _____

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Specific DNA amplification

H.A. Erlich, D.H. Gelfand and R.K. Saiki

*The polymerase chain reaction can synthesize millions of copies of a specific DNA sequence in a brief in vitro reaction. (Thermostable DNA polymerase from *Thermus aquaticus* improves the technique.)*

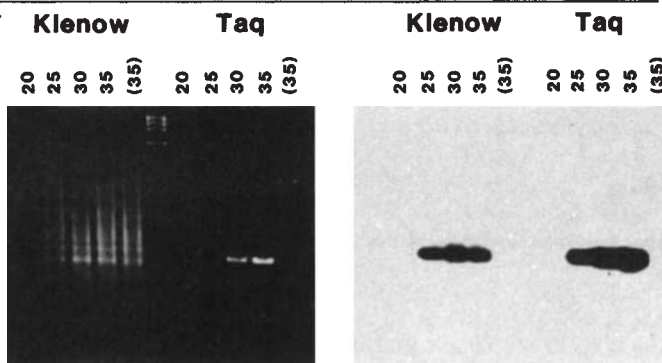
THE analysis of specific nucleotide sequences has become increasingly important in recent years in the clinical diagnosis of genetic and infectious diseases. The polymerase chain reaction (PCR)^{1,2} is an *in vitro* method for the primer-directed enzymatic amplification of specific DNA sequences. This technique is capable of synthesizing over a million copies of a specific target DNA sequence in a few hours, significantly facilitating all subsequent analytic procedures.

The specificity of PCR amplification is based on two oligonucleotide primers that flank the DNA segment to be amplified and hybridize to opposite strands. The procedure involves repeated cycles of heat denaturation of the DNA, annealing of the primers to their complementary sequences, and extension of the annealed primers with DNA polymerase. These primers are oriented so that DNA synthesis by the polymerase proceeds across the region between the primers. The extension products of one primer can serve as a template for the other primer, so each successive cycle essentially doubles the amount of DNA synthesized in the previous cycle. This results in the exponential accumulation of the specific target fragment at approximately 2^n , where n is the number of cycles. The termini of this discrete fragment are defined by the 5' ends of the PCR primers because each of these oligonucleotides becomes physically incorporated into one strand of the PCR product. The efficiency of PCR amplification is not significantly affected by the presence of sequences at the 5' end of the primers that are mismatched with the original template DNA³. This property allows the introduction of additional sequence information, such as restriction site linkers, nucleotide substitutions, deletions, insertions, or regulatory elements, into the PCR product through the oligonucleotide primers.

Thermostable polymerase

The original PCR protocols and most of the early published reports used the Klenow fragment of *Escherichia coli* DNA polymerase I to catalyse the extension of the annealed primers. This required the addition of fresh enzyme during each cycle, however, because the thermolabile Klenow fragment was inactivated by the heat denaturation step used to separate the newly-synthesized strands of DNA. The Klenow-catalysed PCR also yielded a

Fig. 1 Comparison of Klenow and *Taq* PCR performed on human genomic DNA with primers specific for a 110-bp segment of the β -globin gene. Samples were subjected to 20–35 cycles of PCR. *a*, electrophoretic analysis. *b*, Southern analysis with a β -globin probe. Parentheses indicate negative controls with DNA lacking the β -globin gene.



variety of products, only some of which were the target fragment³. Although this did not pose difficulties when the PCR products were analysed with hybridization probes, it limited the use of PCR amplification in applications requiring the production of a specific DNA fragment. To overcome both of these problems, a thermostable DNA polymerase purified from the thermophilic bacterium *Thermus aquaticus* was introduced into the reaction⁴.

This heat-resistant polymerase, termed *Taq* polymerase, is relatively unaffected by the denaturation temperature and does not need to be replenished at each cycle. The modification not only simplifies the procedure, making it amenable to automation, but also substantially increases the specificity of the reaction (Fig. 1), because the annealed primers are extended at 70 °C, near the optimal 70–75 °C temperature range for *Taq* polymerase. The use of *Taq* polymerase significantly increases the overall yield of the PCR technique, and allows the amplification of target DNA sequences up to 2.5 kilobase pairs (from genomic templates)⁴.

The exponential accumulation of PCR-amplified products is not an unlimited process. Eventually, a level of amplification is reached where more primer-template substrate has accumulated than the amount of enzyme present can extend completely in the allotted time. When this occurs, the efficiency of the reaction declines and the amount of PCR product accumulates in a linear rather than an exponential manner. After many cycles, renaturation of the amplified target fragment may also contribute to this "plateau" effect. The increased specificity of the *Taq* polymerase-catalysed reaction over the Klenow-catalysed PCR reduces the competition for the polymerase from non-specific extension products, allowing

more of the target fragment to accumulate before the "plateau" is reached. With *Taq* polymerase-catalysed PCR, the amplification and detection of a single copy of a target DNA sequence in 10 μ g of genomic DNA is possible⁴.

Applications

Although the first published report appeared only two years ago², PCR amplification has already been applied to the diagnosis of genetic disorders^{2,5,7}, the detection of infectious disease pathogens^{8,12}, the study of activated oncogenes^{13,14}, and the analysis of allelic sequence variations¹⁵. In genetic analysis, after a specific locus is amplified, such as β -globin, a variety of techniques can be used to distinguish allelic variants. The presence or absence of an informative restriction site can be determined by using a specific oligonucleotide probe that spans the cleavage site⁶, or, when the amplification reaction results in a unique band, by restriction enzyme digestion of the PCR product and subsequent gel electrophoresis^{1,7}. Short allele-specific oligonucleotide (ASO) probes that hybridize differentially to normal and to mutant or polymorphic sequences represent a more general approach to genetic diagnosis or typing¹. Recently, this method has been applied to the detection of somatic *ras* mutations in peripheral blood cells from acute myelogenous leukemia (AML) patients¹⁴ and in the study of colorectal tumors¹³. The ASO approach has also been used in a dot-blot format with non-radioactive probes to provide a simple, rapid diagnostic test for the hereditary hemoglobinopathies and for the genetic identification of forensic samples, including individual hairs¹⁶. Allelic variation has been determined by using primers specific to conserved sequences to amplify polymorphic regions

and then cloning the PCR products into M13 vectors for sequencing^{3,15}. More recently, this has been accomplished by direct sequencing of the amplified products^{17,18}.

In infectious disease diagnosis, PCR has been used to detect the presence of pathogen-specific sequences from the viruses associated with AIDS (HIV)⁸⁻¹⁰, adult T-cell leukemia (HTLV-I)¹¹, and cervical cancer (HPV)¹². The use of short oligonucleotides based on conserved regions to prime amplification may allow the identification of as-yet uncharacterized sequences related to known groups of viruses¹⁹. The ability to amplify and detect rare pathogen sequences in the presence of a vast excess of host nucleic acids illustrates the broad potential of PCR for clinical diagnosis.

In general, the highly specific nature of *Taq* polymerase-mediated PCR is capable of generating unique fragments from both DNA and RNA templates in yields and purities comparable to fragments prepared from clonally isolated recombinants. These PCR products can be used as substrates for a variety of analytic procedures or as specific hybridization probes. The sensitivity of the procedure should enable the analysis of gene expression or rearrangement in single cells. By virtue of the exponential accumulation of copies derived from a single progenitor DNA sequence, PCR based on *Taq* DNA polymerase represents a form of *in vitro* molecular cloning that can accomplish in an automated 3-4-hour reaction what might otherwise take days or weeks of biological growth and biochemical purification. □

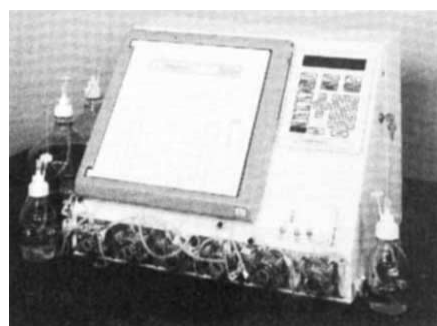
Henry A. Erlich, David H. Gelfand, and Randall K. Saiki are at the Departments of Human Genetics and Microbial Genetics, Cetus Corporation, 1400 Fifty-Third Street, Emeryville, California 94608, USA. For more information, fill in reader service number 100.

- Mullis, K.B. & Faloona, F.A. *Meth Enzym.* **155**, 335-350 (1987).
- Saiki, R.K. *et al.* *Science* **230**, 1350-1354 (1985).
- Scharf, S.J., Horn, G.T. & Erlich, H.A. *Science* **233**, 1076-1078 (1986).
- Saiki, R.K. *et al.* *Science*, **239**, 487-491 (1988).
- Saiki, R.K., Bugawan, T., Horn, G., Mullis, K. & Erlich, H.A. *Nature* **324**, 163-166 (1986).
- Embury, S. *et al.* *New Engl. J. Med.* **316**, 656-661 (1987).
- Kogan, S.C., Doherty, M. & Gitschier, J. *New Engl. J. Med.* **317**, 985-990 (1987).
- Kwok, S. *et al.* *J. Virol.* **61**, 1690-1694 (1987).
- Ou, C.-Y. *et al.* *Science*, **239**, 295-297 (1988).
- Kwok, S. *et al.* (in preparation).
- Duggan, D. *et al.* *Blood* (in the press).
- Shibata, D., Martin, W.J. & Arnheim, N. *J. exp. Med.* **167**, 225-230 (1988).
- Farr, C.J., Saiki, R.K., Erlich, H.A., McCormick, S. & Marshall C.J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Bos, J. *et al.* *Nature* **327**, 293-297 (1987).
- Horn, G.T., Bugawan, T.L., Long, C.M. & Erlich, H.A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Higuchi, R., von Beroldingen, C.H., Sensabaugh, G. & Erlich, H.A. (in preparation).
- Wrischnik, L.A. *et al.* *Nucleic Acids Res.* **15**, 529-542 (1987).
- Wong, C. *et al.* *Nature* **330**, 384-386 (1987).
- Mack, D.H. & Sninsky, J.J. (in preparation).

Biotechnology in Miami

Biotechnologists of every stripe will be travelling to Miami, Florida next week for the Miami Bio/Technology Winter Symposium. A sampling of the products to be on display in the exhibit hall are below.

ENDOTRONICS has an **automated cell culture system** that fulfils middle-range needs, providing product quantities of about 4-10 grams per month — just enough for pre-clinical trials or small-scale production (*Reader Service No. 101*). The \$30,000 (US) microprocessor-controlled Acusyst-Jr. has hollow-fibre perfusion



The Endotronics Acusyst-Jr. cell culture system.

bioreactors that are suited for the continuous production of secreted cell products. Endotronics has overcome the problems of anoxic microenvironments and tapering growth near the end of conventional perfusion bioreactors by the addition of a pressurized expansion chamber to equalize the flow of media throughout the cartridge. Six peristaltic pumps are provided for the addition of growth factors and the harvesting of product. Endotronics will be exhibiting the Acusyst-Jr. in booth 28.

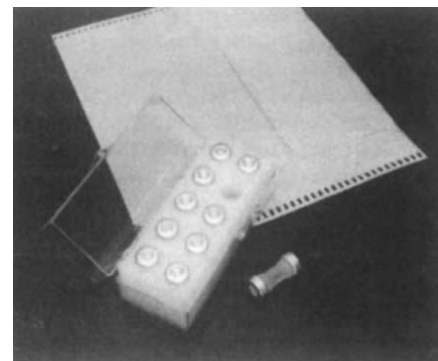
Celltech is drawing upon its expertise in recombinant protein production, gene expression, protein biochemistry, fermentation and downstream processing to offer **research and development services** on a contract basis (*Reader Service No. 102*). Celltech has a new laboratory and process development complex and an established cell culture manufacturing facility for servicing its international customer base. The company also has a range of proprietary technologies for licensing, covering expression and production of recombinant proteins in bacteria, yeast and mammalian cells. Celltech will be in booths 59 and 60.

In booth 11, Cellular Products will be displaying its line of monoclonal antibodies and **growth factors** (*Reader Service No. 103*). Cellular Products produces human B-cell growth factor in 10 ml and 25 ml quantities, free of inducing agent PHA-P, TCGF and immune interferon. The company also sells human platelet-derived growth factor in 100-unit vials, for

the growth of fibroblasts, and human immune interferon produced by T-cells incubated in RPMI 1640 in the presence of human serum albumin and PHA-P. The human immune interferon is sold in 2 ml quantities, containing 250,000 units each.

Peptides and proteins

Applied Biosystems, in booth 39, has expanded its **synthetic peptide service** to include both difficult custom syntheses and a range of off-the-shelf peptides (*Reader Service No. 104*). The company says its list of biologically active synthetic peptides, which includes magainin 1, magainin 2, hANF, rANF and hGRF 1-44, have been proven to be as active as the isolated natural molecules. Each shipment is accompanied by sequence name and composition, HPLC chromatogram, amino acid analysis, and ninhydrin analysis for determination of the coupling efficiency of each synthesis cycle. Options available include the conjugation of protein carriers, FAB mass spectrometry, sequencing, and HF cleavage on peptide



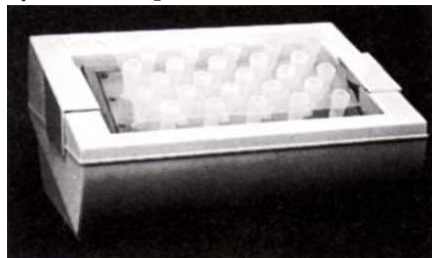
Off-the-shelf peptides from Applied Biosystems.

resins provided by the customer. The synthetic peptides can be ordered in quantities from 1 mg to 1 g, for \$300-500 (US) and \$10,000-20,000 (US) respectively.

Biosym will be demonstrating the new **protein engineering** features of their molecular biology software packages in booth 30 (*Reader Service No. 105*). Discover, Biosym's program for molecular dynamics and simulation for small and macromolecular systems, has been updated to include free energy calculations using the Perturbation Method. The revised program will allow studies of how mutation of a group on a ligand can affect ligand-receptor binding, how reaction rates in enzyme-substrate encounters can be affected by changes in substrate, and how site-directed mutagenesis of a protein affects the thermodynamic stability of its

native conformation. Insight, Biosym's three-dimensional realtime computer graphics program, can provide energy feedback to assist interactive docking of substrates into the active sites of enzymes. Insight can create a contoured energy grid to estimate the interaction energies in real time as the substrate is moved within the active site. The Discover program costs \$100,000 (US) and the Insight program sells for \$25,000 (US). Academicians receive a 90 per cent discount.

The R^MMPS multiple peptide synthesis system from DuPont is a semi-manual system for producing 150 mg each of up to 25 20-mer peptides simultaneously (Reader Service No. 106). The \$1,900 (US) R^MMPS system uses Fmoc chemistry, eliminating the need for HF. Each kit



DuPont's system makes up to 150 mg of peptide.

contains virtually everything needed to synthesize the peptide standard, asn-arg-lys-leu-asp-ile-ala-amide, along with suggested quality control methods. Reagents and resins for five syntheses of the peptide standard are included. The system comes with an instruction manual, 25-peptide R^MMPS processor, five reaction cartridges, amino acids, ninhydrin, control peptide standard and non-routine organic solvents. DuPont will be in booths 4 and 5.

The MilliGen division of Millipore will be introducing its fully automatic peptide synthesizer in booths 8 and 9 (Reader Service No. 107). The MilliGen 960 PepSynthesizer is based on Fmoc-polyamide chemistry and features automatic injection of amino acids, with cycle times of less than 60 minutes per coupling, says MilliGen. No double coupling is required for standard protocols, using the three-column, continuous flow instrument. The



HF cleavage from Multiple Peptide Systems.

instrument's Express-Peptide software performs reagent calculations, controls instrument operation, and provides full documentation of a synthesis. A UV detector allows continuous monitoring of the peptide synthesis process. The PepSynthesizer provides for the unattended coupling of up to 72 amino acids.

Multiple Peptide Systems, in booth 17, will be exhibiting its HF cleavage system, which uses liquified HF in the cleavage of peptide resins and glycoproteins, for use in the synthesis of large amounts of synthetic peptides (Reader Service No. 108). The system's reaction vessels are available in either Kel-F or polypropylene, in a variety of different sizes. The system can cleave up to 24 samples simultaneously or one 100 g sample, and is designed to work with all protocols. The Multiple Peptide Systems HF cleavage system costs \$5,400–6,600 (US).

The PDQuest II analytical software from Protein Databases, Inc. in booth 10 provides both quantitative and positional (*M_r* and *pI*) protein information from 2-dimensional electrophoretic gels (Reader Service No. 109). The software is available for intermediate-level MASSCOMP graphics workstations, and costs \$90,000 (US) for a combined software and hardware package. The software can be used to analyse gels stained by radiolabelling, silver staining or Coomassie techniques, and allows investigators to compare gels and compile databases of gel information. It can perform protein spot matching and quantitative analysis of 3,000 proteins.

Genes and things

The RiboClone cDNA synthesis system from Promega is designed for the synthesis of double-stranded cDNA using poly(A)⁺ mRNA from any source (Reader Service No. 110). The system is available with either the classical oligo(dT) primer, for use when preparing cDNA libraries in vectors where inserts are cloned into a single *EcoRI* site, or with a primer-adaptor that consists of oligo(dT) adjacent to an *XbaI* site, which can be followed by oligo(dT) synthesis and double-digestion with *XbaI* and *EcoRI* to yield molecules with *XbaI* termini on one end and *EcoRI* termini on

the other. Promega will be exhibiting in booth 50.

Beckman Instruments will now be distributing the Designer Genes selection of synthetic genes offered by British Biotechnology Ltd (Reader Service No. 112). The genes are optimized for expression in bacterial or mammalian cell systems and have built-in cutting sites. Latest range additions to the range are genes for epidermal growth factor, insulin-like growth factor 2, streptavidin, horseradish peroxidase, GM-CSF, G-CSF, tumour necrosis factor alpha, interleukin 1 beta, and interleukin 3, 4, 5 and 6. The genes cost between \$3,000 and \$5,000 (US). Beckman will be exhibiting in booths 56–58.

Pharmacia-LKB will be showing off a myriad of products for biotechnology in booths 1-3 (Reader Service No. 112). The LKB VacuGene blotting system uses a low-pressure vacuum pump to transfer nucleic acids from agarose gels to membranes in 15 to 30 minutes. Pharmacia-LKB says the vacuum pump provides transfer conditions for blotting depuri-



The VacuGene blotting system.

nated DNA fragments from 1 kilobase up to chromosome size, with optimal results from genomic DNA digests. Gels of any size up to 18 × 23 cm can be accommodated by the VacuGene unit.

Biotix, in booth 34, sells a line of fully-programmable DNA synthesizers (Reader Service No. 113). The Biotix DNA 100 series instruments are available with a single column, a single column convertible to two columns, and with double columns. All instruments synthesize oligonucleotides at 0.1, 0.5 or 1.0 μM scales with a reagent capacity of 200-mers at the 1.0 μM scale, says Biotix. Interactive software simplifies the input of sequence parameters and calculates and displays the total time required for synthesis. The DNA 100 series instruments are programmed for operation using beta-cyanoethyl phosphoramidite chemistry, but may be programmed for alternate chemistries. Complete systems, including IBM-compatible computer, start at under \$13,000 (US). □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

ADVERTISEMENT

Medwire Div. stocks a wide variety of **TEFLON COATED***

Platinum alloys,
stainless steel,
silver, multistrand
and single strand
available for prompt
shipment. Write for
descriptive brochure
and price lists.



Medwire Div.
Affiliate of Sigmund Cohn Corp.
123 So. Columbus Ave.
Mt. Vernon, N.Y. 10553
914-664-5300
*Reg. T.M. of E.I. DuPont
Reader Service No. 49

Job opportunities in Australia

Richard Pearson

For 200 years Australia has looked overseas for labour, attracting emigrants from throughout the world. As the fanfares sound for its bicentenary, does Australia still hold its attraction for job seekers?

AUSTRALIA has not been isolated from world economic trends and it also traditionally suffers from a high level of unemployment, which last year averaged 7.5 per cent, down from a high of 10.5 per cent three years earlier. Australia also suffers from selected skill shortages, experiencing the same paradoxes as many other developed countries: over-supply of low-skilled people and shortages of the more highly skilled. So what are the employment opportunities for the highly qualified would-be emigrant of the late 1980s?

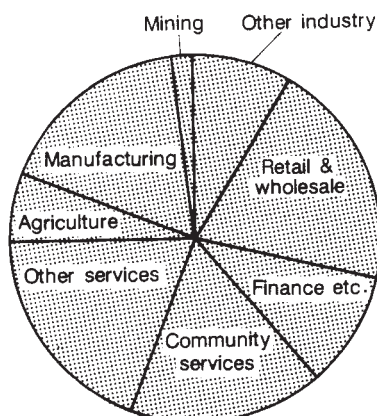
The Australian economy has recovered well from the recession at the start of the decade; indeed its recession lagged behind the rest of the world, reaching its worst in 1982–83, a year or two behind everyone else. Since then, gross domestic product has risen by 4 or 5 per cent annually, although the rate has now fallen back to nearer 3 per cent. Employment has grown strongly over this period, increasing by over 10 per cent since the recession. This growth did not, however, have a commensurate effect on unemployment because of the continuing growth in the size of the labour force and rising participation rates by women. As in Britain, women have been the main beneficiaries of the growth in jobs in the service sector, particularly for part-time workers.

By 1987, services accounted for two out of three jobs, the balance of employment being taken up by manufacturing and other industry, and agriculture (Fig. 1). The fastest growing sectors this decade have been financial, business and property services, fuelled in part by deregulation of the financial markets, and community and welfare services, and retailing. Between them, these three sectors accounted for two-thirds of the job growth. Manufacturing employment, by contrast, has shown little change overall in recent years after a long period of decline. In the mid-1960s, it accounted for one in five jobs; it is now down to one in six.

In line with international experience, the highly qualified have benefited from this employment growth, with job numbers increasing by 5 per cent annually, and skill shortages are starting to reappear. As a consequence, the number of employer-inspired visas for skilled immigrants doubled over the past three years to 3,500 last year.

At an occupational level, the service-based occupations grew fastest following sectoral growth trends, with health and

data-processing leading the way, along with social scientists and skilled service occupations such as motor mechanics and chefs. For scientists, the picture has been mixed. In the physical sciences, where there are about 14,000 professionals, there has been a small surplus. In geoscience and mining-related occupations, a long period of growth came to an end with the recession; but demand has since started picking up again and some regional shortages have been appearing for experienced staff. For chemists, demand generally fluctuates with the economy but currently tends to oversupply, as does demand for physicists, although they are more insu-



Sectoral employment (source: *Labour Force Australia*)

lated from economic fluctuations, having a high dependency on the public sector. Life scientists form a similarly sized group, with over half employed in the public sector, research or scientific institutions. As such, their job prospects are heavily dependent on the state, which has been constraining employment growth resulting in generally limited employment opportunities. The one group of scientists in strong demand are the food scientists, who are mainly employed in the private sector, and who have been benefiting from the growth in the fast-food sector and the rising interest in nutrition and diet.

Computer-related occupations once again appear at the top of the list of good employment prospects, demand having risen spectacularly since the early 1970s. There are now over 25,000 computer programmers and systems analysts in Australia, with more than 500 immigrants entering in each of the past five years. The recession dented the growth in demand but this has turned up again strongly for

experienced staff. Finally, there have been mixed fortunes for the 50,000 engineers with job prospects for many being closely linked to the fluctuating fortunes of the mining and construction industries. The new supply of graduate engineers has also fluctuated markedly over the years in response to both perceived market conditions and changing student interests. A downturn in intakes in the 1970s led to shortages in the early 1980s. Since then, supply has picked up faster than opportunities leading to difficulties for many new graduates; with a contrasting growth in opportunities, there have been shortages of experienced staff. The strongest demand is for electronics, industrial and civil engineers; the lowest for mechanicals.

The outlook for the skilled labour market generally is one of continuing strong demand for occupations linked to the service sector, most notably computing and health professionals, while prospects for other groups will depend on their sector's economic performance. A further complication in the balance between supply and demand has been the erratic trend in output from training schemes and some undergraduate programmes such as engineering. For example, apprentice intakes in the early 1980s fell from 49,000 to 35,000 in a two-year period and the effects of this fall will continue to be felt until the end of the decade.

Australia, like almost every other country, now recognizes the need to invest in developing a highly qualified and skilled workforce if it is to be competitive in world markets. The government has therefore embarked on a strategy of improving training prospects for young people, increasing participation in education after the age of 16 to match the records of countries such as the United States and Japan in secondary education, boosting the number of places in higher education, and encouraging closer co-operation between industry and education. The impact on the more immediate shortages of certain types of highly specialized, experienced staff will, however, be only limited. This is where the country's traditional reliance on immigration is likely to help would-be emigrants, provided they have the highly specialist skills needed. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK.